

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:15:47 ; Search time 32 Seconds
(without alignments)
1010.172 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764
Sequence: 1 MLFVLTCLAVPAISTKS.....DFLLGSSTVAAEKQDPQEA 764

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- Issued Patents AA: *
- 1: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
 - 2: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
 - 3: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
 - 4: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
 - 5: /cgn2_6/prodata/1/aa/PCITUS_COMB.pep:*
 - 6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	746	97.6	746	3 US-08-434-000A-4	Sequence 4, Appl 1
2	746	97.6	746	4 US-09-312-157-4	Sequence 4, Appl 1
3	607	79.5	608	4 US-09-095-385-4	Sequence 4, Appl 1
4	61	8.0	61	3 US-08-856-383-10	Sequence 10, Appl 1
5	61	8.0	61	3 US-08-856-383-11	Sequence 11, Appl 1
6	61	8.0	61	4 US-09-475-088-10	Sequence 10, Appl 1
7	61	8.0	61	4 US-09-475-088-11	Sequence 11, Appl 1
8	60	7.9	60	3 US-08-955-937A-6	Sequence 6, Appl 1
9	60	7.9	60	3 US-09-300-985-6	Sequence 6, Appl 1
10	53	6.9	109	3 US-08-961-564A-9	Sequence 9, Appl 1
11	43	5.6	43	3 US-08-955-937A-11	Sequence 11, Appl 1
12	43	5.6	43	3 US-09-300-985-11	Sequence 11, Appl 1
13	31	4.1	31	3 US-08-856-383-4	Sequence 4, Appl 1
14	31	4.1	31	4 US-09-475-088-4	Sequence 4, Appl 1
15	29	3.8	46	3 US-08-955-937A-10	Sequence 10, Appl 1
16	29	3.8	46	3 US-09-300-985-10	Sequence 10, Appl 1
17	27	3.5	57	3 US-08-955-937A-5	Sequence 5, Appl 1
18	27	3.5	57	3 US-09-300-985-5	Sequence 5, Appl 1
19	21	2.7	757	3 US-08-434-000A-6	Sequence 6, Appl 1
20	21	2.7	757	4 US-09-312-157-6	Sequence 6, Appl 1
21	19	2.5	40	3 US-08-856-383-6	Sequence 6, Appl 1
22	19	2.5	40	4 US-09-475-088-6	Sequence 6, Appl 1
23	17	2.2	769	3 US-08-434-000A-10	Sequence 10, Appl 1
24	17	2.2	769	4 US-09-312-157-10	Sequence 10, Appl 1
25	16	2.1	16	3 US-08-782-480-45	Sequence 45, Appl 1
26	16	2.1	16	3 US-08-954-211-45	Sequence 45, Appl 1
27	16	2.1	16	4 US-09-005-167A-45	Sequence 45, Appl 1

ALIGNMENTS

RESULT 1
US-08-434-000A-4
Sequence 4, Application US/08434000A
Patent No. 6046037
GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,000A
FILING DATE: 12/30/94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/367,395
FILING DATE: 12/30/94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-3510
TELEX: SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
DESCRIPTION: Human Polymunoglobulin Receptor
US-08-434-000A-4

Query Match 97.6%; Score 746; DB 3; Length 746;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIGPEEVNVEGNSVITCYPTSVNHRTRKRYWCROGARGCITLISSEGVSSKY 78
 1 KSPIGPEEVNVEGNSVITCYPTSVNHRTRKRYWCROGARGCITLISSEGVSSKY 60

79 AGRALJTNPENGTGVVNIQAOLSDSDSGRYKCGCLINSRGLSFDVSLSEVSQGPGLNDTK 138
 61 AGRALJTNPENGTGVVNIQAOLSDSDSGRYKCGCLINSRGLSFDVSLSEVSQGPGLNDTK 120

139 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYIDSSGYVNPNTGRIRLDIQGT 198
 121 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYIDSSGYVNPNTGRIRLDIQGT 180

199 GOLLEFSVYNQLRLSDAGQYLCQAGDDSNKKNADLOYLKPPELVEYEDLRSYTFPCA 258
 181 GOLLEFSVYNQLRLSDAGQYLCQAGDDSNKKNADLOYLKPPELVEYEDLRSYTFPCA 240

259 LGPEVANAFAFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPDKDGSFSVYITGLRKE 318
 241 LGPEVANAFAFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPDKDGSFSVYITGLRKE 300

319 DAGRIILCGAHSDGOLOEGSPIOAMOLFVNEESTIPRSPYVKGAVGSSVAVLCPTNRKES 378
 301 DAGRIILCGAHSDGOLOEGSPIOAMOLFVNEESTIPRSPYVKGAVGSSVAVLCPTNRKES 360

379 KSIKYYWCLMEGAONRCPLVDSEGVNKAQYGRSLLEPENGTFVTLNOLTSRDAGF 438
 361 KSIKYYWCLMEGAONRCPLVDSEGVNKAQYGRSLLEPENGTFVTLNOLTSRDAGF 420

439 YWCLTNGDPLMRTTVEIKIIEGEPNLKYVGNVTAVLGETLKYPCHPCKFSSEYKWKCK 498
 421 YWCLTNGDPLMRTTVEIKIIEGEPNLKYVGNVTAVLGETLKYPCHPCKFSSEYKWKCK 480

499 NNTGCOALPDSODEGSKAFVNCDSRLVSLTLNLTADDEGMVKGAVGHPYGETAAV 558
 481 NNTGCOALPDSODEGSKAFVNCDSRLVSLTLNLTADDEGMVKGAVGHPYGETAAV 540

559 YVAVERKKAAGSRDVSILAKADAAPDEKVLDSGFREIENKAIDPRLFAEKAVADTRDA 618
 541 YVAVERKKAAGSRDVSILAKADAAPDEKVLDSGFREIENKAIDPRLFAEKAVADTRDA 600

619 DCSRAVSDGSSEDEGSSRALVSTLPLGLVLAAGAVALGARAHRKNVDVRSIRSYR 678
 601 DCSRAVSDGSSEDEGSSRALVSTLPLGLVLAAGAVALGARAHRKNVDVRSIRSYR 660

679 TDISMSDFENSREFGANDMGAASITQETSLGCKEFTVATTESTTETKEPKAKKSSKEE 738
 661 TDISMSDFENSREFGANDMGAASITQETSLGCKEFTVATTESTTETKEPKAKKSSKEE 720

739 AEMAYKDFLOSSTVAEAODGPOEA 764
 721 AEMAYKDFLOSSTVAEAODGPOEA 746

STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 storage

COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/312,157
 FILING DATE: 14-May-1999
 CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/434,000
 FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Guise, Jeffrey W.
 REGISTRATION NUMBER: 34,613
 REFERENCE/DOCKET NUMBER: 212/127

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 552-8400
 TELEFAX: (619) 552-0159
 TELEX: 67-351

SEQUENCE LISTING
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 746 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: Human Polymunoglobulin Receptor
 US-09-312-157-4

Query Match 97.6%; Score 746; DB 4; Length 746;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIGPEEVNVEGNSVITCYPTSVNHRTRKRYWCROGARGCITLISSEGVSSKY 78
 1 KSPIGPEEVNVEGNSVITCYPTSVNHRTRKRYWCROGARGCITLISSEGVSSKY 60

79 AGRALJTNPENGTGVVNIQAOLSDSDSGRYKCGCLINSRGLSFDVSLSEVSQGPGLNDTK 138
 61 AGRALJTNPENGTGVVNIQAOLSDSDSGRYKCGCLINSRGLSFDVSLSEVSQGPGLNDTK 120

139 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYIDSSGYVNPNTGRIRLDIQGT 198
 121 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYIDSSGYVNPNTGRIRLDIQGT 180

199 GOLLEFSVYNQLRLSDAGQYLCQAGDDSNKKNADLOYLKPPELVEYEDLRSYTFPCA 258
 181 GOLLEFSVYNQLRLSDAGQYLCQAGDDSNKKNADLOYLKPPELVEYEDLRSYTFPCA 240

259 LGPEVANAFAFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPDKDGSFSVYITGLRKE 318
 241 LGPEVANAFAFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPDKDGSFSVYITGLRKE 300

319 DAGRIILCGAHSDGOLOEGSPIOAMOLFVNEESTIPRSPYVKGAVGSSVAVLCPTNRKES 378
 301 DAGRIILCGAHSDGOLOEGSPIOAMOLFVNEESTIPRSPYVKGAVGSSVAVLCPTNRKES 360

379 KSIKYYWCLMEGAONRCPLVDSEGVNKAQYGRSLLEPENGTFVTLNOLTSRDAGF 438
 361 KSIKYYWCLMEGAONRCPLVDSEGVNKAQYGRSLLEPENGTFVTLNOLTSRDAGF 420

439 YWCLTNGDPLMRTTVEIKIIEGEPNLKYVGNVTAVLGETLKYPCHPCKFSSEYKWKCK 498
 421 YWCLTNGDPLMRTTVEIKIIEGEPNLKYVGNVTAVLGETLKYPCHPCKFSSEYKWKCK 480

499 NNTGCOALPDSODEGSKAFVNCDSRLVSLTLNLTADDEGMVKGAVGHPYGETAAV 558
 481 NNTGCOALPDSODEGSKAFVNCDSRLVSLTLNLTADDEGMVKGAVGHPYGETAAV 540

RESULT 2
 US-09-312-157-4
 : Sequence 4, Application US/09312157
 : Patent No. 6303341
 : GENERAL INFORMATION:
 : APPLICANT: ANDREW C. HIATT, JULIAN
 : K.-C. MA, THOMAS LEHNER
 : TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
 : PROTEINS IN PLANTS AND THEIR USES
 : NUMBER OF SEQUENCES: 19
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Lyon & Lyon
 : STREET: 633 West Fifth Street
 : Suite 4700
 : CITY: Los Angeles

Db 481 NNTGCOALPSODEPESKAFVNCDENSRLVSLTLNLTTRADEGWTWCYKQGHFYGETAAV 540
 QY 559 YVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQDPRLFAEERKAVADTRQA 618
 Db 541 YVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQDPRLFAEERKAVADTRQA 600
 QY 619 DGSASVDSGSGSEOGGSSRALVSTLPLGLVAVGAVGAARHRKKNVDRYSIRSYR 678
 Db 601 DGSASVDSGSGSEOGGSSRALVSTLPLGLVAVGAVGAARHRKKNVDRYSIRSYR 660
 QY 679 TDISMDFENSREGANDNMGCASSITQETSLGKEEFYATTESTETETKEPKAKRSSKEE 738
 Db 661 TDISMDFENSREGANDNMGCASSITQETSLGKEEFYATTESTETETKEPKAKRSSKEE 720
 QY 739 AEMAYKDFLLQSTVAEAADGPOEA 764
 Db 721 AEMAYKDFLLQSTVAEAADGPOEA 746

RESULT 3 US-09-095-385-4

Sequence 4, Application US/09095385
 Patent No. 6300104

GENERAL INFORMATION:

APPLICANT: Morrison, Sherie L.
 TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
 TITLE OF INVENTION: BY SINGLE CELLS AND METHODS FOR MAKING AND USING
 NUMBER OF INVENTION: SAME
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
 STREET: 11150 Santa Monica Boulevard, Suite 400
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90025
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/095.385
 FILING DATE: 09-JUN-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/050,969
 FILING DATE: 19-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Canady, Karen S
 REGISTRATION NUMBER: 39,927
 REFERENCE/DOCKET NUMBER: 30435.45USU1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 310 445-1140
 TELEFAX: 310 445-9031
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 608 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-095-385-4

Query Match 79.5%; Score 607; DB 4; Length 608;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFVLTCLAVPAISTKSPFGPEEVNSVEGNSVITCYPTSVNRHTRKYWCQGA 60
 Db 1 MLFVLTCLAVPAISTKSPFGPEEVNSVEGNSVITCYPTSVNRHTRKYWCQGA 60

QY 61 RGCITLISSEGVSSRYAGRANLTNPENGTFFVYNAQLSDDSGRYKCGGLINSGLS 120
 Db 61 RGCITLISSEGVSSRYAGRANLTNPENGTFFVYNAQLSDDSGRYKCGGLINSGLS 120
 QY 121 FDSVLEVSOGPGLNDKRYTVDLGRVTTINCPKTKTNAOKRSKLYQIGLYPVLVYDSS 180
 Db 121 FDSVLEVSOGPGLNDKRYTVDLGRVTTINCPKTKTNAOKRSKLYQIGLYPVLVYDSS 180
 QY 181 GYVNPNTGRIKRLDIOGTGOLLFSVNLRLSPAGQYLCQAGDSSNNKNADLYLKP 240
 Db 181 GYVNPNTGRIKRLDIOGTGOLLFSVNLRLSPAGQYLCQAGDSSNNKNADLYLKP 240
 QY 241 EPELVYEDLRGVTFFHCALGP EVANVAKFLCROSSGECVYVNTLGRAPAEGRILLN 300
 Db 241 EPELVYEDLRGVTFFHCALGP EVANVAKFLCROSSGECVYVNTLGRAPAEGRILLN 300
 QY 301 PODKGSFVSIVTIGLRKEDAGRYLCGASHSDGOLOEGSP IOAMOLFVNEESTIPRSPVVK 360
 Db 301 PODKGSFVSIVTIGLRKEDAGRYLCGASHSDGOLOEGSP IOAMOLFVNEESTIPRSPVVK 360
 QY 361 GVAGSSVAVLCPYNRKESKSIKYWCLMEGAQNGRCPLLVDSGCVKAKQYEGRLSLEEPG 420
 Db 361 GVAGSSVAVLCPYNRKESKSIKYWCLMEGAQNGRCPLLVDSGCVKAKQYEGRLSLEEPG 420
 QY 421 NGFTVYLNOLTSRDAGFYWCLNTGDTLMRTVEIKIIEBPULKVGNVTAVALGETLVK 480
 Db 421 NGFTVYLNOLTSRDAGFYWCLNTGDTLMRTVEIKIIEBPULKVGNVTAVALGETLVK 480
 QY 481 PCHEPKCFSSYEXYKWCNNWNTGCOALPSODEGPKAKVNDENSRLVSLTLNLTTRADEG 540
 Db 481 PCHEPKCFSSYEXYKWCNNWNTGCOALPSODEGPKAKVNDENSRLVSLTLNLTTRADEG 540
 QY 541 WYMGVYKQGHFYGETAAVYVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600
 Db 541 WYMGVYKQGHFYGETAAVYVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600
 QY 601 DPLRLFAE 607
 Db 601 DPLRLFAE 607

RESULT 4

US-08-856-383-10
 Sequence 10, Application US/08856383
 Patent No. 6042833

GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.
 TITLE OF INVENTION: Cellular Internalization of PigR Stalk
 NUMBER OF INVENTION: and Associated Ligands
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/856.383
 FILING DATE: No. 6042833 yet assigned
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/018,958
 FILING DATE: 04-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 02307E-067910PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-856-383-10

Query Match 8.0%; Score 61; DB 3; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

578 ADAAPDEKVLDSGFREIENKAIDPRLFAEERKAVADTRDQADGSRASVDSGSSSEEGCGSS 637
|||||
1 ADAAPDEKVLDSGFREIENKAIDPRLFAEERKAVADTRDQADGSRASVDSGSSSEEGCGSS 60

Oy 638 R 638
Db 61 R 61

RESULT 5

US-08-856-383-11
Sequence 11, Application US/08856383
Patent No. 6042833
GENERAL INFORMATION:
APPLICANT: Mostov, Keith E.
TITLE OF INVENTION: Cellular Internalization of pigR Stalk
TITLE OF INVENTION: and Associated Ligands
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,383
FILING DATE: No. 6042833 yet assigned
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,958
FILING DATE: 04-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 02307E-067910PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-856-383-11

Query Match 8.0%; Score 61; DB 3; Length 61;

Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 578 ADAAPDEKVLDSGFREIENKAIDPRLFAEERKAVADTRDQADGSRASVDSGSSSEEGCGSS 637
|||||
1 ADAAPDEKVLDSGFREIENKAIDPRLFAEERKAVADTRDQADGSRASVDSGSSSEEGCGSS 60

Oy 638 R 638
Db 61 R 61

RESULT 6

US-09-475-088-10
Sequence 10, Application US/09475088
Patent No. 6340743
GENERAL INFORMATION:
APPLICANT: Mostov, Keith E.
TITLE OF INVENTION: Cellular Internalization of pigR Stalk
TITLE OF INVENTION: and Associated Ligands
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/475,088
FILING DATE: 30-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 02307E-067910PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-475-088-10

Query Match 8.0%; Score 61; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 578 ADAAPDEKVLDSGFREIENKAIDPRLFAEERKAVADTRDQADGSRASVDSGSSSEEGCGSS 637
|||||
1 ADAAPDEKVLDSGFREIENKAIDPRLFAEERKAVADTRDQADGSRASVDSGSSSEEGCGSS 60

Oy 638 R 638
Db 61 R 61

RESULT 7

US-09-475-088-11

Sequence 11, Application US/09475088
Patent No. 6340743
GENERAL INFORMATION:
APPLICANT: Mostov, Keith E.
TITLE OF INVENTION: Richman-Eisenstat, Janice
and Associated Ligands
CELLULAR INTERNALIZATION OF PIGR STALK
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/475,088
FILING DATE: 30-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 02307E-067910PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-475-088-11
Query Match
Best Local Similarity 8.0%; Score 61; DB 4; Length 61;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 578 ADADPEKVLDSGFREIENKAIDPRLFAEKRAVADTRDQADGSRASVDSGSGSEGGSS 637
1 ADADPEKVLDSGFREIENKAIDPRLFAEKRAVADTRDQADGSRASVDSGSGSEGGSS 60
QY 638 R 638
DB 61 R 61
RESULT 8
US-08-955-937A-6
Sequence 6, Application US/08955937A
Patent No. 6020161
GENERAL INFORMATION:
APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNEH, ALEMSEGED
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
GENE SUPERFAMILY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
CITY: P. O. BOX 980
STATE: VALLEY FORGE
PA

COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,937A
FILING DATE: 17-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,152
FILING DATE: 19-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-937A-6
Query Match
Best Local Similarity 7.9%; Score 60; DB 3; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 353 PRSPYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 412
1 PRSPYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 60
DB
RESULT 9
US-09-300-985-6
Sequence 6, Application US/09300985A
Patent No. 6232441
GENERAL INFORMATION:
APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND W.
APPLICANT: TRUNEH, ALEMSEGED
APPLICANT: HURLE, MARK ROBERT
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE
FILE REFERENCE: GH-70228-1
CURRENT APPLICATION NUMBER: US/09/300,985A
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: US 08/955,937
EARLIER FILING DATE: 1997-10-22
EARLIER APPLICATION NUMBER: US 60/056,152
EARLIER FILING DATE: 1997-08-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 6
LENGTH: 60
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-300-985-6
Query Match
Best Local Similarity 7.9%; Score 60; DB 3; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 353 PRSPYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 412
1 PRSPYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 60
DB

RESULT 10
US-08-961-564A-9
Sequence 9, Application US/08961564A
Patent No. 6114515
GENERAL INFORMATION:
APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
TITLE OF INVENTION: TRUNEH, ALEMSEGED
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,564A
FILING DATE: 30-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,935
FILING DATE: 25-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70236
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-564A-9
Query Match 6.9%; Score 53; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.9e-46;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 64 CITLISSEGVSSKYAGRANLTFNPENGTFVNIQLSDSDSGRYCGLGINS 116
45 CITLISSEGVSSKYAGRANLTFNPENGTFVNIQLSDSDSGRYCGLGINS 97
RESULT 11
US-08-955-937A-11
Sequence 11, Application US/08955937A
Patent No. 6020161
GENERAL INFORMATION:
APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
TITLE OF INVENTION: TRUNEH, ALEMSEGED
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE

STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,937A
FILING DATE: 17-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,152
FILING DATE: 19-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-937A-11
Query Match 5.6%; Score 43; DB 3; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.8e-36;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 413 LSLLEPBGNGTFVILNLTSDAGFYWCITNGDTLMTPTVEI 455
1 LSLLEPBGNGTFVILNLTSDAGFYWCITNGDTLMTPTVEI 43
RESULT 12
US-09-300-985-11
Sequence 11, Application US/09300985A
Patent No. 6232441
GENERAL INFORMATION:
APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND W.
APPLICANT: TRUNEH, ALEMSEGED
APPLICANT: HUBLE, MARK ROBERT
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE
FILE REFERENCE: GH-70228-1
CURRENT APPLICATION NUMBER: US/09/300,985A
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: US 08/955,937
EARLIER FILING DATE: 1997-10-22
EARLIER APPLICATION NUMBER: US 60/056,152
EARLIER FILING DATE: 1997-08-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 43
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-300-985-11
Query Match 5.6%; Score 43; DB 3; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.8e-36;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 413 LSLLEPBGNGTFVILNLTSDAGFYWCITNGDTLMTPTVEI 455
1 LSLLEPBGNGTFVILNLTSDAGFYWCITNGDTLMTPTVEI 43

Db 1 LSLLEPONGTFTVILNOLTSRDAGFYWCLTNGDTMTTVEI 43

RESULT 13

US-08-856-383-4

Sequence 4, Application US/08856383
Patent No. 6042833

GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.
APPLICANT: Richman-Eisenstat, Janice
TITLE OF INVENTION: Cellular Internalization of p19r stalk
TITLE OF INVENTION: and Associated Ligands
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856,383

FILING DATE: No. 6042833 yet assigned

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/018,958

FILING DATE: 04-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 02307E-067910PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-856-383-4

Query Match 4.18; Score 31; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 4e-24;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 608 EKAVADTRDOAGSRASVDSSSEOGGSSR 638
Db 1 EKAVADTRDOAGSRASVDSSSEOGGSSR 31

RESULT 14

US-09-475-088-4

Sequence 4, Application US/09475088
Patent No. 6340743

GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.
APPLICANT: Richman-Eisenstat, Janice
TITLE OF INVENTION: Cellular Internalization of p19r stalk
TITLE OF INVENTION: and Associated Ligands
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/475,088

FILING DATE: 30-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/856,383

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 02307E-067910PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Query Match 4.18; Score 31; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 4e-24;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 608 EKAVADTRDOAGSRASVDSSSEOGGSSR 638
Db 1 EKAVADTRDOAGSRASVDSSSEOGGSSR 31

RESULT 15

US-08-955-937A-10

Sequence 10, Application US/08955937A
Patent No. 6020161

GENERAL INFORMATION:

APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNH, ALEMSGED
TITLE OF INVENTION: P19r-1, A MEMBER OF IMMUNOGLOBULIN
TITLE OF INVENTION: GENE SUPERFAMILY
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,937A

FILING DATE: 17-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/056,152

FILING DATE: 19-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GH-70228

TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-937A-10

Query Match 3.88; Score 29; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 6.5e-22;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 99 QLSQDDSGRYKCGIGINSRGLSPDVSLEY 127
|||
18 QLSQDDSGRYKCGIGINSRGLSPDVSLEY 46

Search completed: September 9, 2003, 23:21:17
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:18:43 ; Search time 65 Seconds

(Without alignments)
1613.846 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764
Sequence: 1 MLFLVTLCLAVFAIPAIKTSKSP...DFLLQSTVAEAADQGPQEA 764

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 513375 seqs, 137303645 residues

Word size: 0

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	764	100.0	764	10 US-09-818-247-1	Sequence 1, Appli
2	764	100.0	764	10 US-09-881-353-59	Sequence 59, Appl
3	764	100.0	764	10 US-09-989-919-124	Sequence 124, App
4	746	97.6	746	10 US-09-982-107-4	Sequence 4, Appli
5	607	79.5	608	10 US-09-950-294-4	Sequence 4, Appli
6	243	31.8	243	12 US-09-969-748C-108	Sequence 108, Appl
7	205	26.8	602	14 US-10-047-542-51	Sequence 51, Appl
8	95	12.4	95	10 US-09-818-247-20	Sequence 20, Appl
9	94	12.3	102	12 US-09-969-748C-104	Sequence 104, App
10	90	11.8	90	10 US-09-818-247-17	Sequence 17, Appl
11	50	6.5	70	9 US-09-925-299-1218	Sequence 1218, Ap
12	50	6.5	70	11 US-09-925-299-1218	Sequence 1218, Ap
13	47	6.2	244	12 US-09-969-748C-109	Sequence 109, App
14	31	4.1	31	10 US-09-818-247-19	Sequence 19, Appl
15	22	2.9	94	12 US-09-969-748C-107	Sequence 107, App

16	21	2.7	757	10 US-09-818-247-2	Sequence 2, Appl1
17	21	2.7	757	10 US-09-982-107-6	Sequence 6, Appl1
18	20	2.6	102	12 US-09-969-748C-103	Sequence 103, App
19	18	2.4	18	12 US-09-969-748C-71	Sequence 71, Appl
20	18	2.4	23	12 US-09-969-748C-72	Sequence 72, Appl
21	18	2.4	24	12 US-09-969-748C-73	Sequence 73, Appl
22	17	2.2	249	12 US-09-969-748C-110	Sequence 110, App
23	17	2.2	769	10 US-09-818-247-3	Sequence 10, Appl
24	17	2.2	771	12 US-09-982-107-10	Sequence 10, Appl
25	17	2.2	771	12 US-09-982-107-10	Sequence 13, Appl
26	16	2.1	771	10 US-09-818-247-4	Sequence 4, Appl1
27	16	2.1	771	10 US-09-982-107-8	Sequence 8, Appl1
28	13	1.7	732	10 US-09-818-247-5	Sequence 5, Appl1
29	13	1.7	773	10 US-09-818-247-6	Sequence 6, Appl1
30	13	1.7	773	10 US-09-982-107-2	Sequence 2, Appl1
31	11	1.4	258	12 US-09-969-748C-111	Sequence 11, App
32	10	1.3	10	12 US-09-969-748C-74	Sequence 74, Appl
33	9	1.2	9	10 US-09-818-247-12	Sequence 12, Appl
34	9	1.2	624	11 US-09-491-322-22	Sequence 22, Appl
35	8	1.0	106	9 US-09-864-761-43512	Sequence 43512, A
36	8	1.0	140	10 US-09-970-088-11	Sequence 11, Appl
37	8	1.0	140	10 US-09-970-088-13	Sequence 13, Appl
38	8	1.0	140	10 US-09-970-088-14	Sequence 14, Appl
39	8	1.0	777	15 US-10-105-901-3	Sequence 3, Appl1
40	8	1.0	802	9 US-09-758-386-3	Sequence 3, Appl1
41	8	1.0	802	12 US-10-354-358-44	Sequence 44, Appl
42	8	1.0	802	15 US-10-011-548-33	Sequence 33, Appl
43	8	1.0	1298	10 US-09-982-610-33	Sequence 33, Appl
44	8	1.0	1298	15 US-10-262-538-32	Sequence 32, Appl
45	8	1.0	1362	15 US-10-105-901-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-818-247-1
Sequence 1, Application US/09818247
Patent No. US20020102657A1
GENERAL INFORMATION:
APPLICANT: Mostov, Keith E.
APPLICANT: Chaplin, Steven J.
APPLICANT: Richman-Eisenstat, Janice
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Stalk Region of p19r and Methods of Use
FILE REFERENCE: 18062E-000910US
CURRENT APPLICATION NUMBER: US/09/818, 247
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: WO PCT/US01/09699
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,197
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/192,198
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 764
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human polimeric immunoglobulin receptor (p19r)
US-09-818-247-1

Query Match 100.0% Score 764: DB 10: Length 764:
Best Local Similarity 100.0% Pred. NO. 0:
Matches 764: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Oy 1 MLFLVTLCLAVFAIPAIKTSKSPIFGPEEVNSVSGNSVITCYPTSVNRHRTKYRCOGA 60
DB 1 MLFLVTLCLAVFAIPAIKTSKSPIFGPEEVNSVSGNSVITCYPTSVNRHRTKYRCOGA 60

Qy 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAOLSDODSGRYKCGLINSRGLS 120
Db 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAOLSDODSGRYKCGLINSRGLS 120
Qy 121 FDVLSLEVSOQPGLLNLTNTKYTVLDGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180
Db 121 FDVLSLEVSOQPGLLNLTNTKYTVLDGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180
Qy 181 GYVNPVYTGRIRLDIOGTGQLLFVYVINOIRLSDAQOYLCOAGDDSNKKNADLOYLKP 240
Db 181 GYVNPVYTGRIRLDIOGTGQLLFVYVINOIRLSDAQOYLCOAGDDSNKKNADLOYLKP 240
Qy 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAREGRILLN 300
Db 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAREGRILLN 300
Qy 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360
Db 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360
Qy 361 GYAGSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPCG 420
Db 361 GYAGSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPCG 420
Qy 421 NGFTYVILNOLTSRDAGFYWCLTNGDTLMRTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Db 421 NGFTYVILNOLTSRDAGFYWCLTNGDTLMRTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Qy 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540
Db 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540
Qy 541 WYWCYKQOCHFYEETAAVYVAAVEERKAAGSRDVS LAKADAAPDEKYLDSGFREIENKAIO 600
Db 541 WYWCYKQOCHFYEETAAVYVAAVEERKAAGSRDVS LAKADAAPDEKYLDSGFREIENKAIO 600
Qy 601 DPLFAEERKAVADTPOADGSRASVSGSSEBOGSSRALVSTLVPLGLVLAAGAAYGV 660
Db 601 DPLFAEERKAVADTPOADGSRASVSGSSEBOGSSRALVSTLVPLGLVLAAGAAYGV 660
Qy 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASITTOETSLGKEEFVATTE 720
Db 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASITTOETSLGKEEFVATTE 720
Qy 721 STTEKEPKKAKRSSKEEAEMAYKDFLLOSTVAAEAQDGPQEA 764
Db 721 STTEKEPKKAKRSSKEEAEMAYKDFLLOSTVAAEAQDGPQEA 764

RESULT 2
US-09-981-353-59
; Sequence 59, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 59
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 2680109CD1
US-09-981-353-59

Query Match 100.0%; Score 764; DB 10; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLFVLTCLLAFFAISTKSPIFGPEEVNSVEGNSVITCYXPPISVNRHTRKYWCROGA 60
Db 1 MLFVLTCLLAFFAISTKSPIFGPEEVNSVEGNSVITCYXPPISVNRHTRKYWCROGA 60
Qy 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAOLSDODSGRYKCGLINSRGLS 120
Db 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAOLSDODSGRYKCGLINSRGLS 120
Qy 121 FDVLSLEVSOQPGLLNLTNTKYTVLDGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180
Db 121 FDVLSLEVSOQPGLLNLTNTKYTVLDGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180
Qy 181 GYVNPVYTGRIRLDIOGTGQLLFVYVINOIRLSDAQOYLCOAGDDSNKKNADLOYLKP 240
Db 181 GYVNPVYTGRIRLDIOGTGQLLFVYVINOIRLSDAQOYLCOAGDDSNKKNADLOYLKP 240
Qy 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAREGRILLN 300
Db 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAREGRILLN 300
Qy 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360
Db 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360
Qy 361 GYAGSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPCG 420
Db 361 GYAGSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPCG 420
Qy 421 NGFTYVILNOLTSRDAGFYWCLTNGDTLMRTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Db 421 NGFTYVILNOLTSRDAGFYWCLTNGDTLMRTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Qy 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540
Db 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540
Qy 541 WYWCYKQOCHFYEETAAVYVAAVEERKAAGSRDVS LAKADAAPDEKYLDSGFREIENKAIO 600
Db 541 WYWCYKQOCHFYEETAAVYVAAVEERKAAGSRDVS LAKADAAPDEKYLDSGFREIENKAIO 600
Qy 601 DPLFAEERKAVADTPOADGSRASVSGSSEBOGSSRALVSTLVPLGLVLAAGAAYGV 660
Db 601 DPLFAEERKAVADTPOADGSRASVSGSSEBOGSSRALVSTLVPLGLVLAAGAAYGV 660
Qy 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASITTOETSLGKEEFVATTE 720
Db 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASITTOETSLGKEEFVATTE 720
Qy 721 STTEKEPKKAKRSSKEEAEMAYKDFLLOSTVAAEAQDGPQEA 764
Db 721 STTEKEPKKAKRSSKEEAEMAYKDFLLOSTVAAEAQDGPQEA 764

RESULT 3
US-09-989-919-124
; Sequence 124, Application US/09989919
; Patent No. US20020164344A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Pluta, Jason
; APPLICANT: Ghosh, Malavika
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes at
; FILE REFERENCE: DEX-0289
; CURRENT APPLICATION NUMBER: US/09/989.919
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,505
; NUMBER OF SEQ ID NOS: 124

SOFTWARE: PatentIn version 3.1
SEQ ID NO 124
LENGTH: 764
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-919-124

Query Match 100.0%; Score 764; DB 10; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFLVLTCLLAFPAISRTSPFGEENVSVSGNSYSTICYPPTSVNHTTKYKRCGA 60
DB 1 MLFLVLTCLLAFPAISRTSPFGEENVSVSGNSYSTICYPPTSVNHTTKYKRCGA 60
QY 61 RGCCTLLISSEGVSSKYAGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGIGINSRGLS 120
DB 61 RGCCTLLISSEGVSSKYAGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGIGINSRGLS 120
QY 121 FDVSLSEVSGPGLNDTKYVTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSS 180
DB 121 FDVSLSEVSGPGLNDTKYVTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSS 180
QY 181 GYVNPNTGRTIRLDIOGTGQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLP 240
DB 181 GYVNPNTGRTIRLDIOGTGQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLP 240
QY 241 EPELYEDLRGSVTFHCAIGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPAFEGRIILN 300
DB 241 EPELYEDLRGSVTFHCAIGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPAFEGRIILN 300
QY 301 PODKDGFSVYITGLRKEDAGRYLCGASHDQLOEGSPLOAMOLFVNEESTIPRSPTVYK 360
DB 301 PODKDGFSVYITGLRKEDAGRYLCGASHDQLOEGSPLOAMOLFVNEESTIPRSPTVYK 360
QY 361 GVAGSSVAVLCIPYNKESKSTIKYKLMGCAQNGRCPLVDSEGMWKAQYEGRLSLEEPG 420
DB 361 GVAGSSVAVLCIPYNKESKSTIKYKLMGCAQNGRCPLVDSEGMWKAQYEGRLSLEEPG 420
QY 421 NGTFVILNQLTSRDAGFYWCLTNGDTLMTTVEIKIIEGEPNLKYPGNVTAVLGFTLKV 480
DB 421 NGTFVILNQLTSRDAGFYWCLTNGDTLMTTVEIKIIEGEPNLKYPGNVTAVLGFTLKV 480
QY 481 PCHFPCKFSSEYKCYCKNNNTGCCALPQSDGSPKAFVNCDENSRLVSLTLNLTTRADEG 540
DB 481 PCHFPCKFSSEYKCYCKNNNTGCCALPQSDGSPKAFVNCDENSRLVSLTLNLTTRADEG 540
QY 541 WYMGCKGHPGETAAVYVAVEERKAAGSRVSLAKADAPEKVLDSGFREIENKAIQ 600
DB 541 WYMGCKGHPGETAAVYVAVEERKAAGSRVSLAKADAPEKVLDSGFREIENKAIQ 600
QY 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGV 660
DB 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGV 660
QY 661 ARAHRKRVNDVRSIRSYRTDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVATTE 720
DB 661 ARAHRKRVNDVRSIRSYRTDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVATTE 720
QY 721 STTETKEPKKAKRSKKEEAMAYKDFLLQSSSTVAEAODGPOEA 764
DB 721 STTETKEPKKAKRSKKEEAMAYKDFLLQSSSTVAEAODGPOEA 764

RESULT 4
US-09-982-107-4

Sequence 4, Application US/09982107
Patent No. US2002015958A1
GENERAL INFORMATION:
APPLICANT: HIAIT, ANDREM C.
APPLICANT: HEIN, MICH B.
TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING
PROTECTION PROTEINS IN PLANTS AND THEIR USE

FILE REFERENCE: EP13002E
CURRENT APPLICATION NUMBER: US/09/982.107
CURRENT FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 746
TYPE: PRT
ORGANISM: Homo sapiens
US-09-982-107-4

Query Match 97.6%; Score 746; DB 10; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KSPITGPEVNSVEGNSYSTICYPPTSVNHTTKYKRCGARGCITLLISSEGVSSKY 78
DB 1 KSPITGPEVNSVEGNSYSTICYPPTSVNHTTKYKRCGARGCITLLISSEGVSSKY 78
QY 79 AGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGIGINSRGLSFDVSLSEVSGPGLNDTK 138
DB 79 AGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGIGINSRGLSFDVSLSEVSGPGLNDTK 138
QY 139 VYTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSSGYNPNTGRTIRLDIOGT 180
DB 139 VYTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSSGYNPNTGRTIRLDIOGT 180
QY 199 GOLLEFSVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 258
DB 199 GOLLEFSVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 258
QY 240 GOLFESVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 240
DB 240 GOLFESVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 240
QY 259 LGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPAFEGRIILNPODKGFSVYITGLRKE 318
DB 259 LGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPAFEGRIILNPODKGFSVYITGLRKE 318
QY 319 DAGRYLCGASHDQLOEGSPLOAMOLFVNEESTIPRSPTVYKGVAGSSVAVLCIPYNKES 378
DB 319 DAGRYLCGASHDQLOEGSPLOAMOLFVNEESTIPRSPTVYKGVAGSSVAVLCIPYNKES 378
QY 361 DAGRYLCGASHDQLOEGSPLOAMOLFVNEESTIPRSPTVYKGVAGSSVAVLCIPYNKES 360
DB 361 DAGRYLCGASHDQLOEGSPLOAMOLFVNEESTIPRSPTVYKGVAGSSVAVLCIPYNKES 360
QY 379 KSIRKWMLEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 438
DB 379 KSIRKWMLEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 438
QY 439 KSIRKWMLEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 420
DB 439 KSIRKWMLEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 420
QY 441 YWCLTNGDTLMTTVEIKIIEGEPNLKYPGNVTAVLGFTLKPCHFPCKFSSEYKCYCKW 498
DB 441 YWCLTNGDTLMTTVEIKIIEGEPNLKYPGNVTAVLGFTLKPCHFPCKFSSEYKCYCKW 498
QY 499 NNTGQALPQSDGSPKAFVNCDENSRLVSLTLNLTTRADEGMWCKGHPGETAAV 558
DB 499 NNTGQALPQSDGSPKAFVNCDENSRLVSLTLNLTTRADEGMWCKGHPGETAAV 558
QY 541 NNTGQALPQSDGSPKAFVNCDENSRLVSLTLNLTTRADEGMWCKGHPGETAAV 540
DB 541 NNTGQALPQSDGSPKAFVNCDENSRLVSLTLNLTTRADEGMWCKGHPGETAAV 540
QY 559 YVAVEERKAAGSRVSLAKADAPEKVLDSGFREIENKAIQDPLFAEKAADVADTRDQ 618
DB 559 YVAVEERKAAGSRVSLAKADAPEKVLDSGFREIENKAIQDPLFAEKAADVADTRDQ 618
QY 619 DGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGARAHRKRVNDVRSIRSYR 678
DB 619 DGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGARAHRKRVNDVRSIRSYR 678
QY 661 TDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVATTESTETKEPKKAKRSKEE 738
DB 661 TDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVATTESTETKEPKKAKRSKEE 738
QY 739 AEMAYKDFLLQSSSTVAEAODGPOEA 764
DB 739 AEMAYKDFLLQSSSTVAEAODGPOEA 764

RESULT 5
US-09-950-294-4

Sequence 4, Application US/09950294
Patent No. US20020127645A1
GENERAL INFORMATION:

APPLICANT: Morrison, Sherie L.
Chinalacharuvu, Kote R.
TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
BY SINGLE CELLS AND METHODS FOR MAKING AND USING
SAME
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 11150 Santa Monica Boulevard, Suite 400
City: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/950.294
FILING DATE: 10-Sep-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095.385
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Canady, Karen S
REGISTRATION NUMBER: 39,927
REFERENCE/DOCKET NUMBER: 30435.45USU1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310 445-1140
TELEFAX: 310 445-9031
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-950-294-4
Query Match 79.5%; Score 607; DB 10; Length 608;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPLVLTCLLAVFPALSTKSPFSGPEEVNSVEGNSVSYTCYVPTSVNRRTRKRYMCRGA 60
DB 1 MLPLVLTCLLAVFPALSTKSPFSGPEEVNSVEGNSVSYTCYVPTSVNRRTRKRYMCRGA 60
QY 61 RGCITLISSEGVSSKYAGRANLJTNFPENGTFVNIQLSDODSGRYKGLGINSRGLS 120
DB 61 RGCITLISSEGVSSKYAGRANLJTNFPENGTFVNIQLSDODSGRYKGLGINSRGLS 120
QY 121 FDVLSLEVSGCPGLNDTKYTVLGRVTINCPFKTENAKRSLKQIOLYPLVYIDS 180
DB 121 FDVLSLEVSGCPGLNDTKYTVLGRVTINCPFKTENAKRSLKQIOLYPLVYIDS 180
QY 181 GYVNPNTGRIRLDIOGTGOLFSVYINQLRLSDAGGYLCOAGDDSNKKNADLOVLRP 240
DB 181 GYVNPNTGRIRLDIOGTGOLFSVYINQLRLSDAGGYLCOAGDDSNKKNADLOVLRP 240
QY 241 EPELVYEDLRGVTFCALGPPEVANYAKFLCROSSGENDCVVNTLGKRAPAREGRILLN 300
DB 241 EPELVYEDLRGVTFCALGPPEVANYAKFLCROSSGENDCVVNTLGKRAPAREGRILLN 300
QY 301 POKDGSFVYITGLRKEDAGRYLCAHSDGQLOEGSPIAQLFVNEESTIRSPVYVK 360
DB 301 POKDGSFVYITGLRKEDAGRYLCAHSDGQLOEGSPIAQLFVNEESTIRSPVYVK 360
QY 361 GVAGSSVAVLCPNRRKSKSIKYWCLMEGAQNGRCPLVDSEGKVKQYGRISLLEPG 420

DB 361 GVAGSSVAVLCPNRRKSKSIKYWCLMEGAQNGRCPLVDSEGKVKQYGRISLLEPG 420
QY 421 NGTFVYILNOLTSRDAGFYWCLNTGDTLMTTVEIKIIEGPNLKPAGNTAVLGETLKV 480
DB 421 NGTFVYILNOLTSRDAGFYWCLNTGDTLMTTVEIKIIEGPNLKPAGNTAVLGETLKV 480
QY 481 PCHEPCKFSSEYKWKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540
DB 481 PCHEPCKFSSEYKWKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540
QY 541 WYMGVKGQGHFVGETAAVYVAVERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIO 600
DB 541 WYMGVKGQGHFVGETAAVYVAVERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIO 600
QY 601 DPLRLFAE 607
DB 601 DPLRLFAE 607
RESULT 6
US-09-969-748C-108
; Sequence 108, Application US/09969748C
; Publication No. US20030161809A1
; GENERAL INFORMATION:
; APPLICANT: ARIZENE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven
; APPLICANT: BASU, Amresh
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY
; FILE REFERENCE: 057220-0303
; CURRENT APPLICATION NUMBER: US/09/969,748C
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/267,601
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/248,819
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/248,478
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/237,929
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-748C-108
Query Match 31.8%; Score 243; DB 12; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.3e-232;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 492 EKYWKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEGKWKVQGHF 551
DB 1 EKYWKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEGKWKVQGHF 60
QY 552 YGETAAVYVAVERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIO DPLRLFAE KAV 611
DB 61 YGETAAVYVAVERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIO DPLRLFAE KAV 120
QY 612 ADTRDQADGSRASVDSGSEEOGSSRALVSTLPLGLVLAVGAVAVGARARRHRKNVDR 671
DB 121 ADTRDQADGSRASVDSGSEEOGSSRALVSTLPLGLVLAVGAVAVGARARRHRKNVDR 180
QY 672 VSTRSYRTDISMSDFENSRFGANDMKAGSITQETSLGKKEFVATTSSTETTKPKKA 731
DB 181 VSTRSYRTDISMSDFENSRFGANDMKAGSITQETSLGKKEFVATTSSTETTKPKKA 240
QY 732 KRS 734

Db 241 KRS 243

RESULT 7

US-10-047-542-51

Sequence 51, Application US/10047542
 Publication No. US20020168367A1
 GENERAL INFORMATION:
 APPLICANT: LARRICK, JAMES W.
 APPLICANT: WYCOFF, KEITH L.
 TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
 TITLE OF INVENTION: AND BACTERIAL DISEASES
 FILE REFERENCE: 030905.0004.C1P1
 CURRENT APPLICATION NUMBER: US/10/047.542
 CURRENT FILING DATE: 2001-10-26
 PRIOR APPLICATION NUMBER: PCT/US01/13932
 PRIOR FILING DATE: 2001-04-28
 PRIOR APPLICATION NUMBER: 60/200,298
 PRIOR FILING DATE: 2000-04-28
 NUMBER OF SEQ ID NOS: 101
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 51
 LENGTH: 602
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-047-542-51

Query Match 26.8%; Score 205; DB 14; Length 602;

Best Local Similarity 100.0%; Pred. No. 1.6e-194; Mismatches 0; Indels 0; Gaps 0;

Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 TRKYCRQAGARGCGCTITLISSEGYSSKYAGRANLTNPENCTFVNIAQLSQDDSGRYKC 110
 |||||||
 DB 50 TRKYWCQAGARGCGCTITLISSEGYSSKYAGRANLTNPENCTFVNIAQLSQDDSGRYKC 109
 |||||||
 QY 111 GLGINSRGLSFDSLEVYSGEGLLNDTKVTVVDIGRTVINCPTENAKRKSLYQIG 170
 |||||||
 DB 110 GLGINSRGLSFDSLEVYSGEGLLNDTKVTVVDIGRTVINCPTENAKRKSLYQIG 169
 |||||||
 QY 171 LYPPLVYDSSGYVPNTGTRRLDIOGTGOLFVSVINOLRLSPAGYLOAGDSSNSNK 230
 |||||||
 DB 170 LYPPLVYDSSGYVPNTGTRRLDIOGTGOLFVSVINOLRLSPAGYLOAGDSSNSNK 229
 |||||||
 QY 231 KNADLOVLPKPEPELVYEDLRGSYTF 255
 |||||||
 DB 230 KNADLOVLPKPEPELVYEDLRGSYTF 254
 |||||||

RESULT 8

US-09-818-247-20

Sequence 20, Application US/09818247
 Patent No. US20020102657A1
 GENERAL INFORMATION:
 APPLICANT: Mostov, Keith E.
 APPLICANT: Richman, Steven J.
 APPLICANT: Richman, Eistenstat, Janice
 TITLE OF INVENTION: The Regents of the University of California
 TITLE OF INVENTION: Ligands directed to the No. US20020102657A1-Sstalk Region of p19R and Methods of Use Ther
 FILE REFERENCE: 18062E-000910US
 CURRENT APPLICATION NUMBER: US/09/818.247
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: WO PCT/US01/09699
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: US 60/192,197
 PRIOR FILING DATE: 2000-03-27
 PRIOR APPLICATION NUMBER: US 60/192,198
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 20
 LENGTH: 95

TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: portion of
 OTHER INFORMATION: human p19R
 US-09-818-247-20

Query Match

Best Local Similarity 12.4%; Score 95; DB 10; Length 95;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CGVKGHFYGETAAYVAVEERKAAGSRDVSIAKADAAPPEKVIDSGFREIENKAIDPPR 603
 |||||||
 DB 1 CGVKGHFYGETAAYVAVEERKAAGSRDVSIAKADAAPPEKVIDSGFREIENKAIDPPR 60
 |||||||
 QY 604 LFAEKAADTRDQADGSRASVDSGSSSEEGGSSR 638
 |||||||
 DB 61 LFAEKAADTRDQADGSRASVDSGSSSEEGGSSR 95
 |||||||

RESULT 9

US-09-969-748C-104

Sequence 104, Application US/09969748C
 Publication No. US20030161809A1
 GENERAL INFORMATION:
 APPLICANT: ARIZKE, PHARMACEUTICALS, INC.
 APPLICANT: HOUSTON, Lou, L.
 APPLICANT: SHERIDAN, Philip, J.
 APPLICANT: HAWLEY, Stephen
 APPLICANT: GLYNN, Jacqueline, M.
 APPLICANT: CHAPIN, Steven
 APPLICANT: BASU, Amarendra
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY
 TITLE OF INVENTION: AGENTS ACROSS CELLULAR BARRIERS
 FILE REFERENCE: 057220-0303
 CURRENT APPLICATION NUMBER: US/09/969.748C
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: US 60/267,601
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: US 60/248,819
 PRIOR FILING DATE: 2000-11-14
 PRIOR APPLICATION NUMBER: US 60/248,478
 PRIOR FILING DATE: 2000-11-13
 PRIOR APPLICATION NUMBER: US 60/237,929
 PRIOR FILING DATE: 2000-10-02
 NUMBER OF SEQ ID NOS: 115
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 104
 LENGTH: 102
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-969-748C-104

Query Match 12.3%; Score 94; DB 12; Length 102;

Best Local Similarity 100.0%; Pred. No. 4.8e-85; Mismatches 0; Indels 0; Gaps 0;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 GVKQGHYGETAAYVAVEERKAAGSRDVSIAKADAAPPEKVIDSGFREIENKAIDPPR 604
 |||||||
 DB 3 GVKQGHYGETAAYVAVEERKAAGSRDVSIAKADAAPPEKVIDSGFREIENKAIDPPR 62
 |||||||
 QY 605 FAEKAVADTRDQADGSRASVDSGSSSEEGGSSR 638
 |||||||
 DB 63 FAEKAVADTRDQADGSRASVDSGSSSEEGGSSR 96
 |||||||

RESULT 10

US-09-818-247-17

Sequence 17, Application US/09818247
 Patent No. US20020102657A1
 GENERAL INFORMATION:
 APPLICANT: Mostov, Keith E.
 APPLICANT: Chapin, Steven J.

APPLICANT: Richman-Eisenstat, Janice
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Ligands directed to the No. US20020102657A1-Secretory Component,
FILE REFERENCE: No. US20020102657A1-Stalk Region of p19R and Methods of Use Therein
CURRENT APPLICATION NUMBER: US/09/818, 247
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: WO PCT/US01/09699
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,197
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/192,198
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 90
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: portion of
OTHER INFORMATION: human p19R encompassing part of domain 5 and
OTHER INFORMATION: domain 6
US-09-818-247-17

Query Match 11.8%; Score 90; DB 10; Length 90;
Best Local Similarity 100.0%; Pred. No. 4e-81;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 549 GHEYGETAAVYAAVEERKKAAGSDVSLAKADAAPDEKVLDSGFEIENKAIDPRLFAEE 608
DB 1 GHEYGETAAVYAAVEERKKAAGSDVSLAKADAAPDEKVLDSGFEIENKAIDPRLFAEE 60
OY 609 KAAVADTRDQADGSRASVDSGSSSEOGGSSR 638
DB 61 KAAVADTRDQADGSRASVDSGSSSEOGGSSR 90

RESULT 11
US-09-925-299-1218
Sequence 1218, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1218
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (2)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1218

Query Match 6.5%; Score 50; DB 9; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.6e-41;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 219 LCOAGDSDNSNKKNADLOVLKPEPELVEDLRGCVTFHCALGPEVANVAK 268
DB 3 LCOAGDSDNSNKKNADLOVLKPEPELVEDLRGCVTFHCALGPEVANVAK 52

RESULT 12
US-09-925-299-1218
Sequence 1218, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1218
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (2)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1218

Query Match 6.5%; Score 50; DB 11; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.6e-41;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 219 LCOAGDSDNSNKKNADLOVLKPEPELVEDLRGCVTFHCALGPEVANVAK 268
DB 3 LCOAGDSDNSNKKNADLOVLKPEPELVEDLRGCVTFHCALGPEVANVAK 52

RESULT 13
US-09-969-748C-109
Sequence 109, Application US/09969748C
Publication No. US20030161809A1
GENERAL INFORMATION:
APPLICANT: ARIZKE PHARMACEUTICALS, INC.
APPLICANT: HUSTON, Lou, L.
APPLICANT: SHERIDAN, Philip, J.
APPLICANT: HAWLEY, Stephen
APPLICANT: GLYNN, Jacqueline, M.
APPLICANT: CHAPIN, Steven
APPLICANT: BASU, Amareesh
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY
TITLE OF INVENTION: AGENTS ACROSS CELLULAR BARRIERS
FILE REFERENCE: 057220-0303
CURRENT APPLICATION NUMBER: US/09/969,748C
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/267,601
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/248,819
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 60/248,478
PRIOR FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US 60/237,929
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 115
SOFTWARE: PatentIn version 3.0
SEQ ID NO 109
LENGTH: 244
TYPE: PRT
ORGANISM: Simian
US-09-969-748C-109

Query Match 6.2%; Score 47; DB 12; Length 244;
Best Local Similarity 100.0%; Pred. No. 4.9e-38;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 666 RKNVDRVSTIRYRTDISMSDFNSRFGANDKMGASSTQETSLGCK 712

Db 176 RKNDRVRSIRSYRFDISMSDFENSEFGANDMGASSTOETSLOGK 222

RESULT 14

US-09-818-247-19
Sequence 19, Application US/09818247
Patent No. US20020102657A1
GENERAL INFORMATION:
APPLICANT: Mostov, Keith E.
APPLICANT: Chapin, Steven J.
APPLICANT: Richman-Eisenstat, Janice
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Stalk Region of p19R and Methods of Use Therein
FILE REFERENCE: 18062E-000910US
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US/09/818,247
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,197
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/192,198
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 31
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: human p19R
OTHER INFORMATION: stalk
US-09-818-247-19

Query Match 4.1%: Score 31; DB 10; Length 31;

Best Local Similarity 100.0%: Pred. No. 5.6e-23; Indels 0; Gaps 0;

Oy 608 EKAVADTRDQADGSRASVDSGSSSEEGGSSR 638
Db 1 EKAVADTRDQADGSRASVDSGSSSEEGGSSR 31

RESULT 15

US-09-969-748C-107
Sequence 107, Application US/09969748C
Publication No. US20030161809A1
GENERAL INFORMATION:
APPLICANT: ARIZEKE PHARMACEUTICALS, INC.
APPLICANT: HOUSTON, Lou, L.
APPLICANT: SHERIDAN, Philip, J.
APPLICANT: HAWLEY, Stephen
APPLICANT: GLYNN, Jacqueline, M.
APPLICANT: CHAPIN, Steven
APPLICANT: BASU, Amareesh
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE
TITLE OF INVENTION: AGENTS ACROSS CELLULAR BARRIERS
FILE REFERENCE: 057220-0303
CURRENT APPLICATION NUMBER: US/09/969,748C
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/267,601
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/248,819
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 60/248,478
PRIOR FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US 60/237,929
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 115
SOFTWARE: PatentIn version 3.0
SEQ ID NO 107
LENGTH: 94

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-969-748C-107

Query Match 2.9%: Score 22; DB 12; Length 94;

Best Local Similarity 100.0%: Pred. No. 1.3e-13; Indels 0; Gaps 0;

Oy 545 GVKQGHFYGETAAVYVAVEERK 566
Db 3 GVKQGHFYGETAAVYVAVEERK 24

Search completed: September 9, 2003, 23:29:34
Job time : 67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:17:23 : Search time 394 Seconds

(without alignments)
1764.411 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764
Sequence: 1 MLFLVLTCLLAVFPAlSTKS.....DFLLQSTVVAEAQDGPQEA 764

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 5728757 seqs, 909918778 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents-AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	12 US-08-860-037-2	Sequence 2, Appl1

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45		
	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0
	17	17	22	23	24	24	25	25	26	26	27	27	28	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
	US-09-316-412-30	US-09-791-537-92050	US-09-818-247-1	US-09-949-016-6254	US-09-949-039-65	US-09-981-353-59	US-09-989-919-124	US-09-989-919-124	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	
	Sequence 30, Appl1	Sequence 92050, A	Sequence 1, Appl1	Sequence 6254, Appl	Sequence 65, Appl1	Sequence 59, Appl1	Sequence 124, Appl	Sequence 59, Appl1	Sequence 4, Appl1	Sequence 4535, Appl	Sequence 13120, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A	Sequence 8119, A

ALIGNMENTS

RESULT 1
US-08-860-037-2
Sequence 2, Application US/08860037
GENERAL INFORMATION:
APPLICANT: Jan van Oostrum
APPLICANT: Fredericus A.M. Asselbergs
APPLICANT: Markus Gutler
TITLE OF INVENTION: Production of Recombinant Secretary
TITLE OF INVENTION: Component
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Michael Glynn, Novartis Corporation,
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: U.S.
ZIP: 07936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/860,037
FILING DATE: 11 June 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04797
FILING DATE: 12 June 1995
ATTORNEY/AGENT INFORMATION:
NAME: Furman, Diane E.
REGISTRATION NUMBER: 31,104
REFERENCE/DOCKET NUMBER: 4-20259/A/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-277-2311
TELEFAX: 908-277-4009
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-860-037-2

Query Match 100.0%; Score 764; DB 12; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFLVTLCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKYKROGA 60
DB 1 MLFLVTLCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKYKROGA 60
QY 61 RGGCITLISEGVSSEKYGANLTFNFPENGTFVNNIAOLSODDSGRYKGLGINSRGLS 120
DB 61 RGGCITLISEGVSSEKYGANLTFNFPENGTFVNNIAOLSODDSGRYKGLGINSRGLS 120
QY 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRSLYKQIGLYPVLYIDSS 180
DB 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRSLYKQIGLYPVLYIDSS 180
QY 181 GYVNPYTGRIRLDIQGTQQLFSVYINOLRLSDAQYICQAGDSDNSKKNAADQVLPK 240
DB 181 GYVNPYTGRIRLDIQGTQQLFSVYINOLRLSDAQYICQAGDSDNSKKNAADQVLPK 240
QY 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
DB 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
QY 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360
DB 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360
QY 361 GVAAGSSVAVLCPYNNRRESKSIKWCMLMEGAONRCPLLYDSGVMYKAOYEGRLSLEEPC 420
DB 361 GVAAGSSVAVLCPYNNRRESKSIKWCMLMEGAONRCPLLYDSGVMYKAOYEGRLSLEEPC 420
QY 421 NGFTFYILNOLTSRDGFWCLNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480
DB 421 NGFTFYILNOLTSRDGFWCLNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480
QY 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
DB 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
QY 541 WYMGVKGQHFYGETAAYVAVEERKAAGSRDYSLAKADAAPDEKVLDSGFREIEKKAIO 600
DB 541 WYMGVKGQHFYGETAAYVAVEERKAAGSRDYSLAKADAAPDEKVLDSGFREIEKKAIO 600
QY 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLPLGLVLAAGAAYGV 660
DB 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLPLGLVLAAGAAYGV 660
QY 661 ARAHRKKNVDRAVSIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720
DB 661 ARAHRKKNVDRAVSIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720

QY 721 STTETKEPKKAKRSSKEEEMAYKDFLLQSSVTAAEADGPOEA 764
DB 721 STTETKEPKKAKRSSKEEEMAYKDFLLQSSVTAAEADGPOEA 764

RESULT 2
US-09-316-412-30
Sequence 30, Application US/09316412
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: IMMUNOMODULATOR POLYPEPTIDE ZSIG57
FILE REFERENCE: 98-23
CURRENT APPLICATION NUMBER: US/09/316,412
CURRENT FILING DATE: 1999-05-20
EARLIER FILING DATE: 1998-06-18
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 764
TYPE: PRT
ORGANISM: Homo sapiens
US-09-316-412-30

Query Match 100.0%; Score 764; DB 17; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFLVTLCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKYKROGA 60
DB 1 MLFLVTLCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKYKROGA 60
QY 61 RGGCITLISEGVSSEKYGANLTFNFPENGTFVNNIAOLSODDSGRYKGLGINSRGLS 120
DB 61 RGGCITLISEGVSSEKYGANLTFNFPENGTFVNNIAOLSODDSGRYKGLGINSRGLS 120
QY 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRSLYKQIGLYPVLYIDSS 180
DB 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRSLYKQIGLYPVLYIDSS 180
QY 181 GYVNPYTGRIRLDIQGTQQLFSVYINOLRLSDAQYICQAGDSDNSKKNAADQVLPK 240
DB 181 GYVNPYTGRIRLDIQGTQQLFSVYINOLRLSDAQYICQAGDSDNSKKNAADQVLPK 240
QY 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
DB 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
QY 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360
DB 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360
QY 361 GVAAGSSVAVLCPYNNRRESKSIKWCMLMEGAONRCPLLYDSGVMYKAOYEGRLSLEEPC 420
DB 361 GVAAGSSVAVLCPYNNRRESKSIKWCMLMEGAONRCPLLYDSGVMYKAOYEGRLSLEEPC 420
QY 421 NGFTFYILNOLTSRDGFWCLNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480
DB 421 NGFTFYILNOLTSRDGFWCLNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480
QY 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
DB 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
QY 541 WYMGVKGQHFYGETAAYVAVEERKAAGSRDYSLAKADAAPDEKVLDSGFREIEKKAIO 600
DB 541 WYMGVKGQHFYGETAAYVAVEERKAAGSRDYSLAKADAAPDEKVLDSGFREIEKKAIO 600
QY 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLPLGLVLAAGAAYGV 660
DB 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLPLGLVLAAGAAYGV 660
QY 661 ARAHRKKNVDRAVSIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720
DB 661 ARAHRKKNVDRAVSIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720

```

Db      661 AARHRKNDVRSISYRFDISMSOFENSRREGANDNMGASSITQETSLGKKEEFVATTE 720
OY      721 STTETKEPKKARRSSKEEAEMAYKDFLLQSSSTVAEAADGPOEA 764
Db      721 STTETKEPKKARRSSKEEAEMAYKDFLLQSSSTVAEAADGPOEA 764

RESULT 3
US-09-791-537-92050
: Sequence 92050, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Dazez, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
: FILE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 92050
: LENGTH: 764
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-791-537-92050

Query Match      100.0%      Score 764: DB 22: Length 764:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 764: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
Db      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
OY      61 RGGCTTLLSSSEGYSKRYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGLGINSRGLS 120
Db      61 RGGCTTLLSSSEGYSKRYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGLGINSRGLS 120
OY      121 FDSVLEVSQGGPLNDTQVYTVVDLGRVTYVINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
Db      121 FDSVLEVSQGGPLNDTQVYTVVDLGRVTYVINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
OY      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQVLYKP 240
Db      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQVLYKP 240
OY      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCRSSGGENCDVYVNTLGKRAPAFEGRIILN 300
Db      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCRSSGGENCDVYVNTLGKRAPAFEGRIILN 300
OY      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLGESP1QAMQLFVNEESTIPRSPYVVK 360
Db      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLGESP1QAMQLFVNEESTIPRSPYVVK 360
OY      361 GVAASSVAVILCPYRNKESKSIKYCWLMEGAONGRCPLLVDSGHWKAYEGRLSLEBPG 420
Db      361 GVAASSVAVILCPYRNKESKSIKYCWLMEGAONGRCPLLVDSGHWKAYEGRLSLEBPG 420
OY      421 NGTFVILNQLTSRDAGRYWCLTNGDTLMRTTVEIKIIEGEBNLKVPGNVAVLGETLKV 480
Db      421 NGTFVILNQLTSRDAGRYWCLTNGDTLMRTTVEIKIIEGEBNLKVPGNVAVLGETLKV 480
OY      481 PCHEPCKFSSEYKXWCKNNNTGCOALPSODEGSKAFVNCDENSRLVTLNLTVYRADEG 540
Db      481 PCHEPCKFSSEYKXWCKNNNTGCOALPSODEGSKAFVNCDENSRLVTLNLTVYRADEG 540
OY      541 WYMGVKGQGHFYGETAAYVAVEERKAAGSRDVSIAKADAAPEDEKVLDSGFEIENKAIQ 600
Db      541 WYMGVKGQGHFYGETAAYVAVEERKAAGSRDVSIAKADAAPEDEKVLDSGFEIENKAIQ 600
OY      601 DPRLFAEKAVALTRDOADGSRASVDSGSSSEBQGSRALVSTLVPLGLVLAAGVAVGV 660

```

```

Db      601 DPRLFAEKAVALTRDOADGSRASVDSGSSSEBQGSRALVSTLVPLGLVLAAGVAVGV 660
OY      661 AARHRKNDVRSISYRFDISMSOFENSRREGANDNMGASSITQETSLGKKEEFVATTE 720
Db      661 AARHRKNDVRSISYRFDISMSOFENSRREGANDNMGASSITQETSLGKKEEFVATTE 720

RESULT 4
US-09-818-247-1
: Sequence 1, Application US/09818247
: GENERAL INFORMATION:
: APPLICANT: Mostov, Keith E.
: APPLICANT: Chapin, Steven J.
: APPLICANT: Richman-Eisenstat, Janice
: TITLE OF INVENTION: The Regents of the University of California
: TITLE OF INVENTION: Ligands Directed to the Non-Secretory Component,
: FILE REFERENCE: 18062E-000910US
: CURRENT APPLICATION NUMBER: US/09/818,247
: CURRENT FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: WO PCT/US01/09699
: PRIOR FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/192,197
: PRIOR FILING DATE: 2000-03-27
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 764
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: human polymorphic Immunoglobulin receptor (p1gR)
US-09-818-247-1

Query Match      100.0%      Score 764: DB 23: Length 764:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 764: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
Db      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
OY      61 RGGCTTLLSSSEGYSKRYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGLGINSRGLS 120
Db      61 RGGCTTLLSSSEGYSKRYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGLGINSRGLS 120
OY      121 FDSVLEVSQGGPLNDTQVYTVVDLGRVTYVINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
Db      121 FDSVLEVSQGGPLNDTQVYTVVDLGRVTYVINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
OY      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQVLYKP 240
Db      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQVLYKP 240
OY      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCRSSGGENCDVYVNTLGKRAPAFEGRIILN 300
Db      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCRSSGGENCDVYVNTLGKRAPAFEGRIILN 300
OY      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLGESP1QAMQLFVNEESTIPRSPYVVK 360
Db      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLGESP1QAMQLFVNEESTIPRSPYVVK 360
OY      361 GVAASSVAVILCPYRNKESKSIKYCWLMEGAONGRCPLLVDSGHWKAYEGRLSLEBPG 420
Db      361 GVAASSVAVILCPYRNKESKSIKYCWLMEGAONGRCPLLVDSGHWKAYEGRLSLEBPG 420

```

OY 421 NGFTVILNOLTSRDAGFWCLNGLDMLRTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
DB 421 NGFTVILNOLTSRDAGFWCLNGLDMLRTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
OY 481 PCHEPKCFSSYEKWKMMNTGQALPDSDEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
DB 481 PCHEPKCFSSYEKWKMMNTGQALPDSDEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
OY 541 WYMGVYKQGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600
DB 541 WYMGVYKQGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600
OY 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660
DB 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660
OY 661 ARAHRKKNDRVIRSIRYRTDISMSDENSREFGANDMGASITTOETSLGKKEEFVATTE 720
DB 661 ARAHRKKNDRVIRSIRYRTDISMSDENSREFGANDMGASITTOETSLGKKEEFVATTE 720
OY 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAEAADGPOEA 764
DB 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAEAADGPOEA 764

RESULT 5
US-09-949-016-6254
; Sequence 6254, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 6254
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6254

Query Match 100.0%; Score 764; DB 24; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKXWROGA 60
DB 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKXWROGA 60
OY 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVNNIAQLSODDSGRYKCGLGINSRGLS 120
DB 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVNNIAQLSODDSGRYKCGLGINSRGLS 120
OY 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVINCPEKTEENAKRSKSLKQIGLFPVLIDSS 180
DB 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVINCPEKTEENAKRSKSLKQIGLFPVLIDSS 180
OY 181 GYVNPNTYGRIRLDIGTGLFSSVYINQLRLSDAGQYLCQAGDSDSNKKNADLOVLKP 240
DB 181 GYVNPNTYGRIRLDIGTGLFSSVYINQLRLSDAGQYLCQAGDSDSNKKNADLOVLKP 240
OY 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGENDCVVNTLGKRAPAFEGRIILN 300
DB 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGENDCVVNTLGKRAPAFEGRIILN 300

OY 301 PODKGSFVSIVITGLRKEDAGRYLCAHSQDLOESPPIOANOLFVNEESTIPRSPTVVK 360
DB 301 PODKGSFVSIVITGLRKEDAGRYLCAHSQDLOESPPIOANOLFVNEESTIPRSPTVVK 360
OY 361 GVAGSSVAALCPYNNRESKSIKWKCLMEGAONRCPLVDSSEGVAQAQYEGRLSLLEERG 420
DB 361 GVAGSSVAALCPYNNRESKSIKWKCLMEGAONRCPLVDSSEGVAQAQYEGRLSLLEERG 420
OY 421 NGFTVILNOLTSRDAGFWCLNGLDMLRTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
DB 421 NGFTVILNOLTSRDAGFWCLNGLDMLRTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
OY 481 PCHEPKCFSSYEKWKMMNTGQALPDSDEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
DB 481 PCHEPKCFSSYEKWKMMNTGQALPDSDEGSKAFVNCDENSRLVSLTLNLYTRADEG 540
OY 541 WYMGVYKQGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600
DB 541 WYMGVYKQGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600
OY 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660
DB 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660
OY 661 ARAHRKKNDRVIRSIRYRTDISMSDENSREFGANDMGASITTOETSLGKKEEFVATTE 720
DB 661 ARAHRKKNDRVIRSIRYRTDISMSDENSREFGANDMGASITTOETSLGKKEEFVATTE 720
OY 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAEAADGPOEA 764
DB 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAEAADGPOEA 764

RESULT 6
US-09-949-039-65
; Sequence 65, Application US/09949039
; GENERAL INFORMATION:
; APPLICANT: HAWLEY, STEPHEN B.
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE
; TITLE OF INVENTION: BINDING REGIONS DIRECTED TO TRANSCRIPTOTIC LIGANDS
; FILE REFERENCE: 057220/1301
; CURRENT APPLICATION NUMBER: US/09/949, 039
; CURRENT FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-039-65

Query Match 100.0%; Score 764; DB 24; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKXWROGA 60
DB 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKXWROGA 60
OY 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVNNIAQLSODDSGRYKCGLGINSRGLS 120
DB 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVNNIAQLSODDSGRYKCGLGINSRGLS 120
OY 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVINCPEKTEENAKRSKSLKQIGLFPVLIDSS 180
DB 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVINCPEKTEENAKRSKSLKQIGLFPVLIDSS 180
OY 181 GYVNPNTYGRIRLDIGTGLFSSVYINQLRLSDAGQYLCQAGDSDSNKKNADLOVLKP 240
DB 181 GYVNPNTYGRIRLDIGTGLFSSVYINQLRLSDAGQYLCQAGDSDSNKKNADLOVLKP 240
OY 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGENDCVVNTLGKRAPAFEGRIILN 300
DB 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGENDCVVNTLGKRAPAFEGRIILN 300

Db 241 EPELVYEDLRGSVTFHCLGPEVANNAFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
Qy 301 PQRDGSFSVITGLRKEDAGRYLCGAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360
Db 301 PQRDGSFSVITGLRKEDAGRYLCGAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360
Qy 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVVKAQYEGRLSLEBPG 420
Db 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVVKAQYEGRLSLEBPG 420
Qy 421 NGFTVILNOLTSRDAGFYWCLTNGDLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKV 480
Db 421 NGFTVILNOLTSRDAGFYWCLTNGDLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKV 480
Qy 481 PCHEPKFSSEYKWKNNNTGCOALPSODEGPKAFVNCDENSRLVSLTLNLTTRADEG 540
Db 481 PCHEPKFSSEYKWKNNNTGCOALPSODEGPKAFVNCDENSRLVSLTLNLTTRADEG 540
Qy 541 WYWCVKOGHFYGETAAYVAVERKKAAGSDVSLAKADAAPEKVLDSGFREIENKA1Q 600
Db 541 WYWCVKOGHFYGETAAYVAVERKKAAGSDVSLAKADAAPEKVLDSGFREIENKA1Q 600
Qy 601 DPLFAEKAADRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVAVGAVAVG 660
Db 601 DPLFAEKAADRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVAVGAVAVG 660
Qy 661 ARAHRKKNVDVRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720
Db 661 ARAHRKKNVDVRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720
Qy 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764
Db 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764

RESULT 7
US-09-981-353-59
: Sequence 59, Application US/09981353
: GENERAL INFORMATION:
: APPLICANT: Lasek, Amy W.
: APPLICANT: Jones, David A.
: TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
: FILE REFERENCE: PA-0038 US
: CURRENT APPLICATION NUMBER: US/09/981.353
: CURRENT FILING DATE: 2001-10-11
: NUMBER OF SEQ ID NOS: 194
: SOFTWARE: PERL Program
: SEQ ID NO 59
: LENGTH: 764
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 2680109CD1
US-09-981-353-59

Query Match 100.0%; Score 764; DB 25; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHRTKRYWCROGA 60
Db 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHRTKRYWCROGA 60
Qy 61 RGGCTTLISSEGYVSSKAGRANLTNPENCTFVYVNTIAQLSDODSGRKCGLGINSRGLS 120
Db 61 RGGCTTLISSEGYVSSKAGRANLTNPENCTFVYVNTIAQLSDODSGRKCGLGINSRGLS 120
Qy 121 FDSVLEVSOGPGLNDTKVTVYVDLGRVTINCPFKTEAOKRKSILYKQIGLYPVLIDSS 180
Db 121 FDSVLEVSOGPGLNDTKVTVYVDLGRVTINCPFKTEAOKRKSILYKQIGLYPVLIDSS 180
Qy 181 GYVNPNTYGRIRLDIOGTGQLLFVSVINQLRLSDAGQYLCAGDDSNSNKNADLOVLKP 240

Db 181 GYVNPNTYGRIRLDIOGTGQLLFVSVINQLRLSDAGQYLCAGDDSNSNKNADLOVLKP 240
Qy 241 EPELVYEDLRGSVTFHCLGPEVANNAFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
Db 241 EPELVYEDLRGSVTFHCLGPEVANNAFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300
Qy 301 PQRDGSFSVITGLRKEDAGRYLCGAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360
Db 301 PQRDGSFSVITGLRKEDAGRYLCGAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360
Qy 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVVKAQYEGRLSLEBPG 420
Db 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVVKAQYEGRLSLEBPG 420
Qy 421 NGFTVILNOLTSRDAGFYWCLTNGDLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKV 480
Db 421 NGFTVILNOLTSRDAGFYWCLTNGDLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKV 480
Qy 481 PCHEPKFSSEYKWKNNNTGCOALPSODEGPKAFVNCDENSRLVSLTLNLTTRADEG 540
Db 481 PCHEPKFSSEYKWKNNNTGCOALPSODEGPKAFVNCDENSRLVSLTLNLTTRADEG 540
Qy 541 WYWCVKOGHFYGETAAYVAVERKKAAGSDVSLAKADAAPEKVLDSGFREIENKA1Q 600
Db 541 WYWCVKOGHFYGETAAYVAVERKKAAGSDVSLAKADAAPEKVLDSGFREIENKA1Q 600
Qy 601 DPLFAEKAADRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVAVGAVAVG 660
Db 601 DPLFAEKAADRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVAVGAVAVG 660
Qy 661 ARAHRKKNVDVRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720
Db 661 ARAHRKKNVDVRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720
Qy 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764
Db 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764

RESULT 8
US-09-989-919-124
: Sequence 124, Application US/09989919
: GENERAL INFORMATION:
: APPLICANT: Macina, Roberto
: APPLICANT: Recipon, Herve
: APPLICANT: Pluta, Jason
: APPLICANT: Ghosh, Malaivika
: APPLICANT: Sun, Yongming
: APPLICANT: Liu, Chenghua
: TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes
: FILE REFERENCE: DEX-0289
: CURRENT APPLICATION NUMBER: US/09/989.919
: CURRENT FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: 60/252.505
: NUMBER OF SEQ ID NOS: 124
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 124
: LENGTH: 764
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-989-919-124

Query Match 100.0%; Score 764; DB 25; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHRTKRYWCROGA 60
Db 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHRTKRYWCROGA 60
Qy 61 RGGCTTLISSEGYVSSKAGRANLTNPENCTFVYVNTIAQLSDODSGRKCGLGINSRGLS 120

```

Db      61 RGCITLISSEGVSSKYAGRAVLTNFPENGTFFVNIQAOLSDODSGRYKCGILNSGLS 120
Qy      121 FDVLEVSOGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLYKQIGLYPVLYDSS 180
Db      121 FDVLEVSOGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLYKQIGLYPVLYDSS 180
Qy      181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQYLCQAGDDSNKKNAADLOVLRP 240
Db      181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQYLCQAGDDSNKKNAADLOVLRP 240
Qy      241 EPELYVEDLRGSYTFHCALGPEVANYAKFLCROSSGENDVYVNTLGRAPAFEGRIILN 300
Db      241 EPELYVEDLRGSYTFHCALGPEVANYAKFLCROSSGENDVYVNTLGRAPAFEGRIILN 300
Qy      301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIQAMOLFVNEESTIPRSPYVK 360
Db      301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIQAMOLFVNEESTIPRSPYVK 360
Qy      361 GVAGSSVAVLCPPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNKAQYEGRLSLLEBP 420
Db      361 GVAGSSVAVLCPPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNKAQYEGRLSLLEBP 420
Qy      421 NGFTVILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Db      421 NGFTVILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Qy      481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLVTADGC 540
Db      481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLVTADGC 540
Qy      541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVSIAKADAAPDEKVLDSGFREIENKAIO 600
Db      541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVSIAKADAAPDEKVLDSGFREIENKAIO 600
Qy      601 DPLFAEKAADVADTROADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVAGV 660
Db      601 DPLFAEKAADVADTROADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVAGV 660
Qy      661 ARAHRKKNDRVIRSTRYDIMSDFENSRFGANDMGASITTOETSLGKEEFVATTE 720
Db      661 ARAHRKKNDRVIRSTRYDIMSDFENSRFGANDMGASITTOETSLGKEEFVATTE 720
Qy      721 STTETKEPKKAKRSSKEEAMAYKDFLOSSTVAAEADGPOEA 764
Db      721 STTETKEPKKAKRSSKEEAMAYKDFLOSSTVAAEADGPOEA 764

```

RESULT 9

```

US-60-239-841-59
: Sequence 59, Application US/60239841
: GENERAL INFORMATION:
: APPLICANT: Lasek, Amy W.
: TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
: FILE REFERENCE: PA-0038 P
: CURRENT APPLICATION NUMBER: US/60/239, 841
: NUMBER OF SEQ ID NOS: 194
: SOFTWARE: PERL Program
: SEQ ID NO 59
: LENGTH: 764
: TYPE: PRP
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: 2680109CD1
US-60-239-841-59

```

```

Query Match      100.0%; Score 764; DB 32; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLLEFVTLCLAAVPAISTKSPIEGPEVANSVEGNSITCYPTPTSVNHRTRKMYCPOGA 60
Db      1 MLLEFVTLCLAAVPAISTKSPIEGPEVANSVEGNSITCYPTPTSVNHRTRKMYCPOGA 60
Qy      61 RGCITLISSEGVSSKYAGRAVLTNFPENGTFFVNIQAOLSDODSGRYKCGILNSGLS 120
Db      61 RGCITLISSEGVSSKYAGRAVLTNFPENGTFFVNIQAOLSDODSGRYKCGILNSGLS 120
Qy      121 FDVLEVSOGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLYKQIGLYPVLYDSS 180
Db      121 FDVLEVSOGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLYKQIGLYPVLYDSS 180
Qy      181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQYLCQAGDDSNKKNAADLOVLRP 240
Db      181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQYLCQAGDDSNKKNAADLOVLRP 240
Qy      241 EPELYVEDLRGSYTFHCALGPEVANYAKFLCROSSGENDVYVNTLGRAPAFEGRIILN 300
Db      241 EPELYVEDLRGSYTFHCALGPEVANYAKFLCROSSGENDVYVNTLGRAPAFEGRIILN 300
Qy      301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIQAMOLFVNEESTIPRSPYVK 360
Db      301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIQAMOLFVNEESTIPRSPYVK 360
Qy      361 GVAGSSVAVLCPPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNKAQYEGRLSLLEBP 420
Db      361 GVAGSSVAVLCPPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNKAQYEGRLSLLEBP 420
Qy      421 NGFTVILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Db      421 NGFTVILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Qy      481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLVTADGC 540
Db      481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTNLVTADGC 540
Qy      541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVSIAKADAAPDEKVLDSGFREIENKAIO 600
Db      541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVSIAKADAAPDEKVLDSGFREIENKAIO 600
Qy      601 DPLFAEKAADVADTROADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVAGV 660
Db      601 DPLFAEKAADVADTROADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVAGV 660
Qy      661 ARAHRKKNDRVIRSTRYDIMSDFENSRFGANDMGASITTOETSLGKEEFVATTE 720
Db      661 ARAHRKKNDRVIRSTRYDIMSDFENSRFGANDMGASITTOETSLGKEEFVATTE 720
Qy      721 STTETKEPKKAKRSSKEEAMAYKDFLOSSTVAAEADGPOEA 764
Db      721 STTETKEPKKAKRSSKEEAMAYKDFLOSSTVAAEADGPOEA 764

```

RESULT 10

```

US-09-717-888-4
: Sequence 4, Application US/09717888
: GENERAL INFORMATION:
: APPLICANT: ANDREW C. HIATT, JULIAN
: TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

```

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/717.888
FILING DATE: 20-Nov-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/434.000
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/367.395
FILING DATE: 30-Dec-94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-3510
SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
DESCRIPTION: Human Polymunoglobulin Receptor
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-717-888-4

Query Match          97.6%: Score 746; DB 21; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KSPFEGEEVNSVSGNSVITCYPPPTSVNRHTRKRYKRCGARGCGITLLISSEGYSSKY 78
DB 1 KSPFEGEEVNSVSGNSVITCYPPPTSVNRHTRKRYKRCGARGCGITLLISSEGYSSKY 60
QY 79 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 138
DB 61 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 120
QY 139 VYTVDLGRVTYINCPEFTENAOKRSLYKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 198
DB 121 VYTVDLGRVTYINCPEFTENAOKRSLYKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 180
QY 199 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 258
DB 181 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 240
QY 259 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 318
DB 241 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 300
QY 319 DAGRYLCGAHSDGOLQSGSPLOAMQLFVNEESTIPRSPPTYVKGAGSSVAVLCPYNRKES 378
DB 301 DAGRYLCGAHSDGOLQSGSPLOAMQLFVNEESTIPRSPPTYVKGAGSSVAVLCPYNRKES 360
QY 379 KSIRYKWLMEGAONGRCPLVLDSEGNVKAQYEGRLSLLEBPGNCTFVILNQLTSRDAGF 438
DB 361 KSIRYKWLMEGAONGRCPLVLDSEGNVKAQYEGRLSLLEBPGNCTFVILNQLTSRDAGF 420
QY 439 YWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKVPCHPFCKSSYEEKYCKW 498
DB 421 YWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKVPCHPFCKSSYEEKYCKW 480
QY 499 NNTGCOALPSODEGSPSAFVNCDENSRLVSLTLNLVTRADEGMYWCGVKGQGHFYGETAAV 558
DB 481 NNTGCOALPSODEGSPSAFVNCDENSRLVSLTLNLVTRADEGMYWCGVKGQGHFYGETAAV 540
QY 559 YVAVEERKAAGSRDVSILAKADAAPDEKVLDSGREIENKAIODPRLFAEKKAVADTRDOA 618
DB 541 YVAVEERKAAGSRDVSILAKADAAPDEKVLDSGREIENKAIODPRLFAEKKAVADTRDOA 600

```

```

DB 541 YVAVEERKAAGSRDVSILAKADAAPDEKVLDSGREIENKAIODPRLFAEKKAVADTRDOA 600
QY 619 DGSRAVSDGSSSEOGGSSRALVSTVPLGLVLAAGAAGVARRRRKRVDRYSIRSYR 678
DB 601 DGSRAVSDGSSSEOGGSSRALVSTVPLGLVLAAGAAGVARRRRKRVDRYSIRSYR 660
QY 679 TDISMDFENSREPGANDNNAGASSITQETSLGKKEEVATTESTETKEKKAKRSKEE 738
DB 661 TDISMDFENSREPGANDNNAGASSITQETSLGKKEEVATTESTETKEKKAKRSKEE 720
QY 739 AEMAYKDFLQSSVFAEAADGPQEA 764
DB 721 AEMAYKDFLQSSVFAEAADGPQEA 746

RESULT 11
US-09-982-107-4
Sequence 4: Application US/09982107
GENERAL INFORMATION:
APPLICANT: HIATT, ANDREW C.
APPLICANT: HEIN, MICHAEL B.
TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING
TITLE OF INVENTION: PROTECTION PROTEINS IN PLANTS AND THEIR USE
FILE REFERENCE: EP13002E
CURRENT APPLICATION NUMBER: US/09/982.107
CURRENT FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 4
LENGTH: 746
TYPE: PRT
ORGANISM: Homo sapiens
US-09-982-107-4

Query Match          97.6%: Score 746; DB 25; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KSPFEGEEVNSVSGNSVITCYPPPTSVNRHTRKRYKRCGARGCGITLLISSEGYSSKY 78
DB 1 KSPFEGEEVNSVSGNSVITCYPPPTSVNRHTRKRYKRCGARGCGITLLISSEGYSSKY 60
QY 79 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 138
DB 61 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 120
QY 139 VYTVDLGRVTYINCPEFTENAOKRSLYKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 198
DB 121 VYTVDLGRVTYINCPEFTENAOKRSLYKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 180
QY 199 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 258
DB 181 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 240
QY 259 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 318
DB 241 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 300
QY 319 DAGRYLCGAHSDGOLQSGSPLOAMQLFVNEESTIPRSPPTYVKGAGSSVAVLCPYNRKES 378
DB 301 DAGRYLCGAHSDGOLQSGSPLOAMQLFVNEESTIPRSPPTYVKGAGSSVAVLCPYNRKES 360
QY 379 KSIRYKWLMEGAONGRCPLVLDSEGNVKAQYEGRLSLLEBPGNCTFVILNQLTSRDAGF 438
DB 361 KSIRYKWLMEGAONGRCPLVLDSEGNVKAQYEGRLSLLEBPGNCTFVILNQLTSRDAGF 420
QY 439 YWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKVPCHPFCKSSYEEKYCKW 498
DB 421 YWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKVPCHPFCKSSYEEKYCKW 480
QY 499 NNTGCOALPSODEGSPSAFVNCDENSRLVSLTLNLVTRADEGMYWCGVKGQGHFYGETAAV 558
DB 481 NNTGCOALPSODEGSPSAFVNCDENSRLVSLTLNLVTRADEGMYWCGVKGQGHFYGETAAV 540

```

```

RESULT 12
US-60-443-566-4535
: Sequence 4535, Application US/60443566
:
: GENERAL INFORMATION:
:
: APPLICANT: CARCILL, Michele
:
: APPLICANT: BEGOVICH, Ann
:
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
:
: TITLE OF INVENTION: RHUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
:
: FILE REFERENCE: CL001447
:
: CURRENT APPLICATION NUMBER: US/60/443,566
:
: CURRENT FILING DATE: 2003-01-30
:
: NUMBER OF SEQ. ID NOS.: 25102
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ. ID NO. 4535
:
: LENGTH: 764
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: US-60-443-566-4535

```

Query Match	86.8%	Score 663	DB 32	Length 764
Best Local Similarity	99.9%	Pred. No. 0		
Matches 763	Conservative	0	Mismatches	1
			Indels	0
			Gaps	0
QY 1 MLLEFVTLCLLAVFAIPAISTKSPIFGPEEVNSVEGNSVSYICYPPPTSVNHRTRKRYCROGA 60	1 MLLEFVTLCLLAVFAIPAISTKSPIFGPEEVNSVEGNSVSYICYPPPTSVNHRTRKRYCROGA 60			
Db 1 MLLEFVTLCLLAVFAIPAISTKSPIFGPEEVNSVEGNSVSYICYPPPTSVNHRTRKRYCROGA 60				
QY 61 RGGCITLISEGCVSSKYGARALTNFPENGTFVYVNIADLSDDDSRKCGIGINSRGIS 120	61 RGGCITLISEGCVSSKYGARALTNFPENGTFVYVNIADLSDDDSRKCGIGINSRGIS 120			
Db 61 RGGCITLISEGCVSSKYGARALTNFPENGTFVYVNIADLSDDDSRKCGIGINSRGIS 120				
QY 121 FVLSLEVSQGPGLLNTTKYTYVDLGRVTYVINGPFTKENQKRKSLYKQIGLCPVLVIDSS 180	121 FVLSLEVSQGPGLLNTTKYTYVDLGRVTYVINGPFTKENQKRKSLYKQIGLCPVLVIDSS 180			
Db 121 FVLSLEVSQGPGLLNTTKYTYVDLGRVTYVINGPFTKENQKRKSLYKQIGLCPVLVIDSS 180				
QY 181 GYVNPYNTGRIRLIDIDGTCGLLFVYVINOILRLSDAQGYLCOAGDDSNSKKKADIQVLYKP 240	181 GYVNPYNTGRIRLIDIDGTCGLLFVYVINOILRLSDAQGYLCOAGDDSNSKKKADIQVLYKP 240			
Db 181 GYVNPYNTGRIRLIDIDGTCGLLFVYVINOILRLSDAQGYLCOAGDDSNSKKKADIQVLYKP 240				
QY 241 EPELTVYEDLRGSVTFHICALGPEVANYAKFLCROSSGENDVYVNTLTKRAPAFEGIRILLN 300	241 EPELTVYEDLRGSVTFHICALGPEVANYAKFLCROSSGENDVYVNTLTKRAPAFEGIRILLN 300			
Db 241 EPELTVYEDLRGSVTFHICALGPEVANYAKFLCROSSGENDVYVNTLTKRAPAFEGIRILLN 300				
QY 301 PDKDGSFSVYITGLKKEDAGRYLGCASHDGOLOEGSPIQAMOLFVNEBSTIPRSPTVYK 360	301 PDKDGSFSVYITGLKKEDAGRYLGCASHDGOLOEGSPIQAMOLFVNEBSTIPRSPTVYK 360			
Db 301 PDKDGSFSVYITGLKKEDAGRYLGCASHDGOLOEGSPIQAMOLFVNEBSTIPRSPTVYK 360				
QY 361 GYVGSVVAVLCPYNRKRESKSIKYMCLMEASONGRCPLVDSEGWYKAQYEGRLSLLEEPG 420	361 GYVGSVVAVLCPYNRKRESKSIKYMCLMEASONGRCPLVDSEGWYKAQYEGRLSLLEEPG 420			
Db 361 GYVGSVVAVLCPYNRKRESKSIKYMCLMEASONGRCPLVDSEGWYKAQYEGRLSLLEEPG 420				
QY 421 NGFTFYILNOLTSRDAGFYWCLTNGDTLARTTVEIKIIEGEPNLKPGNVTAVLCEPLKY 480	421 NGFTFYILNOLTSRDAGFYWCLTNGDTLARTTVEIKIIEGEPNLKPGNVTAVLCEPLKY 480			
Db 421 NGFTFYILNOLTSRDAGFYWCLTNGDTLARTTVEIKIIEGEPNLKPGNVTAVLCEPLKY 480				

```

RESULT 13
US-60-452-680-13120
: Sequence 13120, Application US/60452680
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: GRUPE, Andrew
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001450
: CURRENT APPLICATION NUMBER: US/60/452,680
: CURRENT FILING DATE: 2003-03-07
: NUMBER OF SEQ ID NOS: 116213
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 13120
: LENGTH: 764
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-452-680-13120

```

Query Match	Best Local Match	Similarity	86.8%	Score 663	DB 32	Length 764			
Matches	763	Conservative	0	Mismatches	1	Indels	0	Gaps	0
QY	1	MLEFLVTCLLAVFPAISTKSPIFGPEEYNSVEGNSVITCYYPPTSVNHRTRKYCROGA	60						
Db	1	MLEFLVTCLLAVFPAISTKSPIFGPEEYNSVEGNSVITCYYPPTSVNHRTRKYCROGA	60						
QY	61	RGCGITLISSSGGYSSKTAGRANLTPNPENGFFVYVNIQOLSDSDSGRYKCGIGISRGIS	120						
Db	61	RGCGITLISSSGGYSSKTAGRANLTPNPENGFFVYVNIQOLSDSDSGRYKCGIGISRGIS	120						
QY	121	FDVSELEVSQCGGLNDTRKYTVTVDLGRTYVINCPEKTEAOKRKSILYKQIGLPLVYVIDSS	180						
Db	121	FDVSELEVSQCGGLNDTRKYTVTVDLGRTYVINCPEKTEAOKRKSILYKQIGLPLVYVIDSS	180						
QY	181	GYVNPNTYGRIRLDIQGTGOLFFSVYINQLRLSDAGQYLCQAGDSSNSKKKADQLYLP	240						
Db	181	GYVNPNTYGRIRLDIQGTGOLFFSVYINQLRLSDAGQYLCQAGDSSNSKKKADQLYLP	240						
QY	241	EPELVYEDLRGSYTFHCHALGPEVANVAKFLCRSSGSENCDDVYVNTLGRAPAFEGIRILN	300						
Db	241	EPELVYEDLRGSYTFHCHALGPEVANVAKFLCRSSGSENCDDVYVNTLGRAPAFEGIRILN	300						
QY	301	PODKRGSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIQAMQLFVNEESTIPRSPYVVK	360						
Db	301	PODKRGSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIQAMQLFVNEESTIPRSPYVVK	360						
QY	361	GVASSVAVALCPYNNRKSSTIKYVNCIMWGAONGRCPLLYDSEGWYKAYEGRLSILEBPG	420						
Db	361	GVAGGSAVALCPYNNRKSSTIKYVNCIMWGAONGRCPLLYDSEGWYKAYEGRLSILEBPG	420						
QY	421	NGTFTVILINQLTSDRAGEFWCLTNGDILMRTVTEIKIIEGEPNLKVPNGNVAVLGETLKV	480						

```

Db      |||||||
421  NCTFTVILNQLTSRDAFYWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
Oy      |||||||
481  PCHEPPCKFSSYEKWKCKNNNTGCOALPSQDEGPKSAFVNCDENSRLVSLTINTLYTRADEG 540
Db      |||||||
481  PCHEPPCKFSSYEKWKCKNNNTGCOALPSQDEGPKSAFVNCDENSRLVSLTINTLYTRADEG 540
Oy      |||||||
541  WYMGVGVGKHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGPREIENKAIQ 600
Db      |||||||
541  WYMGVGVGKHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGPREIENKAIQ 600
Oy      |||||||
601  DPLFAEERKAVADTRDOAGDSRASVDSGSSSEEGSSRALVSTLVPGLVLAAGAVAVG 660
Db      |||||||
601  DPLFAEERKAVADTRDOAGDSRASVDSGSSSEEGSSRALVSTLVPGLVLAAGAVAVG 660
Oy      |||||||
661  ARARHRRKVVDRVSTIRSTRDTISMSDFENSREFGANDNNGASSITQETSLGKEEFVATTE 720
Db      |||||||
661  ARARHRRKVVDRVSTIRSTRDTISMSDFENSREFGANDNNGASSITQETSLGKEEFVATTE 720
Oy      |||||||
721  STTEKPEKKAARSSKEEAEMAYKDFLLQSSVVAEAODGPQEA 764
Db      |||||||
721  STTEKPEKKAARSSKEEAEMAYKDFLLQSSVVAEAODGPQEA 764

```

RESULT 14 US-09-949-016-8119

```

: Sequence 8119, Application US/09949016
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8119
: LENGTH: 816
: TYPE: PRT
: ORGANISM: Human
: US-09-949-016-8119

```

Query Match 86.8%; Score 663; DB 24; Length 816;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy      |||||||
1  MLLEPVLCLLAVFAPISTKSPFGEPEVNSVEGNSVSTCYPPPTSVMRHRKRYWCRGA 60
Db      |||||||
53  MLLEPVLCLLAVFAPISTKSPFGEPEVNSVEGNSVSTCYPPPTSVMRHRKRYWCRGA 112
Oy      |||||||
61  RGGCTTLISSEGEVSSKAGRANLNFENGTFFVNIQALSODDSGRKCGIGINSRLS 120
Db      |||||||
113  RGGCTTLISSEGEVSSKAGRANLNFENGTFFVNIQALSODDSGRKCGIGINSRLS 172
Oy      |||||||
121  FDSVLEVSQGPGLNDTKVYTVVDLGRVTINCPEKTENAQRKSLYKOIGLYPVLVIDSS 180
Db      |||||||
173  FDSVLEVSQGPGLNDTKVYTVVDLGRVTINCPEKTENAQRKSLYKOIGLYPVLVIDSS 232
Oy      |||||||
181  GYVNPNTYGRIRLDIOGQGLFVSVINQLRLSDAGQYLCOAGDDSSNNKKRADLOVLPK 240
Db      |||||||
233  GYVNPNTYGRIRLDIOGQGLFVSVINQLRLSDAGQYLCOAGDDSSNNKKRADLOVLPK 292
Oy      |||||||
241  EPELVYEDLRGCVTFHACALGPEVANAKFLCROSSGENDVYVNTLGKRAPAFEGRIILLN 300
Db      |||||||
293  EPELVYEDLRGCVTFHACALGPEVANAKFLCROSSGENDVYVNTLGKRAPAFEGRIILLN 352
Oy      |||||||
301  PODKDGFSFSVYITGLRKEDAGRYLCGANSQQLQEGSPIQAWOLFVNEESTIPRSPYVVK 360

```

```

Db      |||||||
353  PODKDGFSFSVYITGLRKEDAGRYLCGANSQQLQEGSPIQAWOLFVNEESTIPRSPYVVK 412
Oy      |||||||
361  GVAGSSVAIVLCPIYRKSSKSTIKYWCLEMGAGNCGCPLLYOSEGVVKKOYEGRLSLLEPG 420
Db      |||||||
413  GVAGSSVAIVLCPIYRKSSKSTIKYWCLEMGAGNCGCPLLYOSEGVVKKOYEGRLSLLEPG 472
Oy      |||||||
421  NCTFTVILNQLTSRDAFYWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
Db      |||||||
473  NCTFTVILNQLTSRDAFYWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKV 532
Oy      |||||||
481  PCHEPPCKFSSYEKWKCKNNNTGCOALPSQDEGPKSAFVNCDENSRLVSLTINTLYTRADEG 540
Db      |||||||
533  PCHEPPCKFSSYEKWKCKNNNTGCOALPSQDEGPKSAFVNCDENSRLVSLTINTLYTRADEG 592
Oy      |||||||
541  WYMGVGVGKHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGPREIENKAIQ 600
Db      |||||||
593  WYMGVGVGKHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGPREIENKAIQ 652
Oy      |||||||
601  DPLFAEERKAVADTRDOAGDSRASVDSGSSSEEGSSRALVSTLVPGLVLAAGAVAVG 660
Db      |||||||
653  DPLFAEERKAVADTRDOAGDSRASVDSGSSSEEGSSRALVSTLVPGLVLAAGAVAVG 712
Oy      |||||||
661  ARARHRRKVVDRVSTIRSTRDTISMSDFENSREFGANDNNGASSITQETSLGKEEFVATTE 720
Db      |||||||
713  ARARHRRKVVDRVSTIRSTRDTISMSDFENSREFGANDNNGASSITQETSLGKEEFVATTE 772
Oy      |||||||
721  STTEKPEKKAARSSKEEAEMAYKDFLLQSSVVAEAODGPQEA 764
Db      |||||||
773  STTEKPEKKAARSSKEEAEMAYKDFLLQSSVVAEAODGPQEA 816

```

RESULT 15 US-09-950-294-4

```

: Sequence 4, Application US/09950294
: GENERAL INFORMATION:
: APPLICANT: Morison, Sherie L.
: TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
: BY SINGLE CELLS AND METHODS FOR MAKING AND USING
: SAME
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESS: Merchant, Gould, Smith, Edell, Welter & Schmidt
: STREET: 11150 Santa Monica Boulevard, Suite 400
: CITY: Los Angeles
: STATE: CA
: COUNTRY: USA
: ZIP: 90025

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/950,294
FILING DATE: 10-Sep-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,385
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Canada, Karen S
REGISTRATION NUMBER: 39,927
REFERENCE/DOCKET NUMBER: 30435,45USU1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310 445-1140
TELEFAX: 310 445-9031
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-950-294-4

Query Match

79.5%; Score 607; DB 25; Length 608;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 MLLEFVLTCLLAVPAISTKSPIRGPEEVNSVEGNSVSYTCYPTSVNRHTRKYWCROGA 60
        |||
Db      1 MLLEFVLTCLLAVPAISTKSPIRGPEEVNSVEGNSVSYTCYPTSVNRHTRKYWCROGA 60
QY      61 RGGCITLISSEGYSSKRYAGRANLTNFPENGTFVYVNI AQLSODDSGRYKGLGINSRGLS 120
        |||
Db      61 RGGCITLISSEGYSSKRYAGRANLTNFPENGTFVYVNI AQLSODDSGRYKGLGINSRGLS 120
QY      121 FDVLSLEVSGGPGLLNDTKYTVTDLGRTVTINCPFKTENAKRSLYKQIGLYPVLYDSS 180
        |||
Db      121 FDVLSLEVSGGPGLLNDTKYTVTDLGRTVTINCPFKTENAKRSLYKQIGLYPVLYDSS 180
QY      181 GYVNPVNTGRIRLDIOGTGQLFYSVYINOLRLSDAGQYLCQAGDSDSNKKNADLOYLKP 240
        |||
Db      181 GYVNPVNTGRIRLDIOGTGQLFYSVYINOLRLSDAGQYLCQAGDSDSNKKNADLOYLKP 240
QY      241 EPELVYEDLRGSYTFHCALGPEVANYAKFLCRQSSGENCDVYVNTLGKRAPAFEGRIILN 300
        |||
Db      241 EPELVYEDLRGSYTFHCALGPEVANYAKFLCRQSSGENCDVYVNTLGKRAPAFEGRIILN 300
QY      301 PODKDSFSYVITGLRKEDAGRYLGCASHSDGQLQEGSPIOAWQLFVNEESTIPRSPTVVK 360
        |||
Db      301 PODKDSFSYVITGLRKEDAGRYLGCASHSDGQLQEGSPIOAWQLFVNEESTIPRSPTVVK 360
QY      361 GYAGSSVAVYLCPIYNRKESKSIKYWCLMEGAONGRCPLVDSGEMVKAQYEGRLSLEEPG 420
        |||
Db      361 GYAGSSVAVYLCPIYNRKESKSIKYWCLMEGAONGRCPLVDSGEMVKAQYEGRLSLEEPG 420
QY      421 NGFTFYILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
        |||
Db      421 NGFTFYILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
QY      481 PCHFPCFKFSSEYKWKCKMNTGCOALPSODEGSPSKAFVNCDENSRLVSLTLNLVTRADEG 540
        |||
Db      481 PCHFPCFKFSSEYKWKCKMNTGCOALPSODEGSPSKAFVNCDENSRLVSLTLNLVTRADEG 540
QY      541 WYMGVYKQGHFYGETAAVYVAVEERKAGSRDYSIAKADAAPDEKYLDSGFREIENKAIQ 600
        |||
Db      541 WYMGVYKQGHFYGETAAVYVAVEERKAGSRDYSIAKADAAPDEKYLDSGFREIENKAIQ 600
QY      601 DPLRLFAE 607
        |||
Db      601 DPLRLFAE 607
```

Search completed: September 9, 2003, 23:28:01
Job time : 398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:17:53; Search time 15 Seconds

(without alignments)
486.619 Million cell updates/sec

Title: US-09-818-247-1

Sequence: 764

Sequence: 1 MLFVLTCLAVPAISTKS.....DFLLQSTVAEAADGPOEA 764

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 47124 seqs, 9554045 residues

Word size: 0

Total number of hits satisfying chosen parameters: 47124

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Pending Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being predicted,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	2.1	16	US-09-005-318F-45	Sequence 45, Appl
2	16	2.1	16	US-10-062-467B-45	Sequence 45, Appl
3	8	1.0	762	PCT-US02-29560A-285	Sequence 285, App
4	8	1.0	802	PCT-US02-29560A-284	Sequence 284, App
5	8	1.0	1308	US-60-487-610-1446	Sequence 1446, Ap
6	8	1.0	1324	US-60-487-610-1445	Sequence 1445, Ap
7	8	1.0	1379	US-60-487-610-1444	Sequence 1444, Ap
8	7	0.9	21	US-60-490-419-86	Sequence 86, Appl
9	7	0.9	21	US-60-490-149-86	Sequence 86, Appl
10	7	0.9	209	US-60-495-114-2042	Sequence 2042, Ap
11	7	0.9	216	PCT-US03-25276-32	Sequence 32, Appl
12	7	0.9	287	PCT-US02-18638A-176	Sequence 176, App
13	7	0.9	300	PCT-US02-18638A-174	Sequence 174, App
14	7	0.9	300	PCT-US03-10783-2	Sequence 2, Appl
15	7	0.9	300	US-60-487-610-2568	Sequence 2568, App
16	7	0.9	314	PCT-US02-18638A-172	Sequence 172, Appl
17	7	0.9	314	US-09-981-845-1	Sequence 1, Appl
18	7	0.9	314	US-60-487-610-2567	Sequence 2567, Ap
19	7	0.9	336	US-09-897-516A-7391	Sequence 7391, Ap
20	7	0.9	404	PCT-US03-21559-7	Sequence 7, Appl
21	7	0.9	479	US-60-478-196-3256	Sequence 3256, App
22	7	0.9	603	PCT-US02-18638A-186	Sequence 186, App
23	7	0.9	603	US-60-490-890-1440	Sequence 1440, App
24	7	0.9	630	PCT-US02-29560A-364	Sequence 364, App
25	7	0.9	653	US-60-485-450-1067	Sequence 1067, App
26	7	0.9	687	US-60-485-450-1066	Sequence 1066, App

27	7	0.9	851	US-09-876-773-2	Sequence 2, Appl
28	7	0.9	851	US-60-487-610-2240	Sequence 2240, Ap
29	7	0.9	851	US-60-485-450-1426	Sequence 1426, Ap
30	7	0.9	866	PCT-US02-37650-6	Sequence 6, Appl
31	7	0.9	898	US-60-478-196-3332	Sequence 3332, Ap
32	7	0.9	1331	PCT-US03-22467-2	Sequence 2, Appl
33	7	0.9	1333	PCT-US03-22467-12	Sequence 12, Appl
34	7	0.9	1798	US-60-487-610-2811	Sequence 2811, Ap
35	7	0.9	1798	US-60-487-610-2812	Sequence 2812, Ap
36	7	0.9	1954	PCT-US03-19027-2	Sequence 2, Appl
37	7	0.9	3714	US-60-487-610-3753	Sequence 2753, Ap
38	6	0.8	6	US-09-005-318F-41	Sequence 41, Appl
39	6	0.8	6	US-10-062-467B-41	Sequence 41, Appl
40	6	0.8	20	PCT-US03-11766-5	Sequence 5, Appl
41	6	0.8	34	PCT-US03-18657-3	Sequence 3, Appl
42	6	0.8	36	PCT-US03-18657-2	Sequence 2, Appl
43	6	0.8	62	US-10-644-807-251	Sequence 251, App
44	6	0.8	62	US-10-644-807-345	Sequence 345, App
45	6	0.8	82	PCT-US03-25418-27	Sequence 27, Appl

ALIGNMENTS

```
RESULT 1
US-09-005-318F-45
; Sequence 45, Application US/09005318F
; GENERAL INFORMATION:
; APPLICANT: HEIN, MICH B.
; APPLICANT: HINAT, ANDREW C.
; TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
; FILE REFERENCE: EPI3004B
; CURRENT APPLICATION NUMBER: US/09/005.318F
; PRIOR FILING DATE: 1998-01-09
; PRIOR APPLICATION NUMBER: 09/782,481
; PRIOR FILING DATE: 1997-01-10
; PRIOR APPLICATION NUMBER: 09/005,167
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-005-318F-45

Query Match      2.1% Score 16; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 5; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

QY      598 AIDPRLFAEKAVAD 613
DB      1 AIDPRLFAEKAVAD 16
|||||

RESULT 2
US-10-062-467B-45
; Sequence 45, Application US/10062467B
; GENERAL INFORMATION:
; APPLICANT: HEIN, MICH B.
; APPLICANT: HINAT, ANDREW C.
; TITLE OF INVENTION: J CHAIN POLYPEPTIDE TARGETING MOLECULE LINKED TO AN IMAGING
; FILE REFERENCE: EPI3003C
; CURRENT APPLICATION NUMBER: US/10/062.467B
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 08/782,480
; PRIOR FILING DATE: 1997-01-10
; PRIOR APPLICATION NUMBER: 09/005,167
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 45
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-10-062-467B-45

Query Match 2.18; Score 16; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 AIDPPRLFAEKAAYD 613
Db 1 AIDPPRLFAEKAAYD 16

RESULT 3
PCT-US02-29560A-285
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Gish, Kurt C.
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Wilson, Keith E.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-002710PC
CURRENT APPLICATION NUMBER: PCT/US02/29560A
CURRENT FILING DATE: 2002-09-17
PRIOR APPLICATION NUMBER: US 60/323,469
PRIOR FILING DATE: 2001-09-17
NUMBER OF SEQ ID NOS: 412
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 285
LENGTH: 762
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-29560A-285

Query Match 1.0%; Score 8; DB 1; Length 762;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
Db 94 EDAGRYLC 101

RESULT 4
PCT-US02-29560A-284
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Gish, Kurt C.
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Wilson, Keith E.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-002710PC
CURRENT APPLICATION NUMBER: PCT/US02/29560A
CURRENT FILING DATE: 2002-09-17
PRIOR APPLICATION NUMBER: US 60/323,469
PRIOR FILING DATE: 2001-09-17
NUMBER OF SEQ ID NOS: 412
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 284

LENGTH: 802
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-29560A-284

Query Match 1.0%; Score 8; DB 1; Length 802;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
Db 94 EDAGRYLC 101

RESULT 5
US-60-487-610-1446
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: HUANG, Hongjin
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
FILE REFERENCE: CL001469
CURRENT APPLICATION NUMBER: US/60/487,610
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 97101
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1446
LENGTH: 1308
TYPE: PRT
ORGANISM: Homo sapiens
US-60-487-610-1446

Query Match 1.0%; Score 8; DB 7; Length 1308;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
Db 751 EDAGRYLC 758

RESULT 6
US-60-487-610-1445
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: HUANG, Hongjin
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
FILE REFERENCE: CL001469
CURRENT APPLICATION NUMBER: US/60/487,610
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 97101
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1445
LENGTH: 1324
TYPE: PRT
ORGANISM: Homo sapiens
US-60-487-610-1445

Query Match 1.0%; Score 8; DB 7; Length 1324;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
Db 757 EDAGRYLC 764

RESULT 7

US-60-487-610-1444
 ; Sequence 1444, Application US/60487610
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: HUANG, Hongjin
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
 ; FILE REFERENCE: CL001469
 ; CURRENT APPLICATION NUMBER: US/60/487,610
 ; CURRENT FILING DATE: 2003-07-17
 ; NUMBER OF SEQ ID NOS: 97101
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1444
 ; LENGTH: 1379
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-60-487-610-1444

Query Match
 Best Local Similarity 100.0%; Score 8; DB 7; Length 1379;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
 |||||
 Db 757 EDAGRYLC 764

RESULT 8
 US-60-490-419-86
 ; Sequence 86, Application US/60490419
 ; GENERAL INFORMATION:
 ; APPLICANT: Fred E. Regneir
 ; APPLICANT: Jiri Adamec
 ; TITLE OF INVENTION: Digital Chromatography
 ; FILE REFERENCE: 302840.3000-100
 ; CURRENT APPLICATION NUMBER: US/60/490,419
 ; CURRENT FILING DATE: 2003-07-25
 ; NUMBER OF SEQ ID NOS: 277
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 86
 ; LENGTH: 21
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-60-490-419-86

Query Match
 Best Local Similarity 100.0%; Score 7; DB 7; Length 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 705 QETSIGG 711
 |||||
 Db 8 QETSIGG 14

RESULT 9
 US-60-490-149-86
 ; Sequence 86, Application US/60490149
 ; GENERAL INFORMATION:
 ; APPLICANT: Fred E. Regneir
 ; APPLICANT: Jiri Adamec
 ; APPLICANT: Zhang Zhang
 ; TITLE OF INVENTION: Digital Chromatography
 ; FILE REFERENCE: 302840.3000-100
 ; CURRENT APPLICATION NUMBER: US/60/490,149
 ; CURRENT FILING DATE: 2003-07-25
 ; NUMBER OF SEQ ID NOS: 277
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 86
 ; LENGTH: 21
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens

US-60-490-149-86
 Query Match
 Best Local Similarity 100.0%; Score 7; DB 7; Length 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 705 QETSIGG 711
 |||||
 Db 8 QETSIGG 14

RESULT 10
 US-60-495-114-2042
 ; Sequence 2042, Application US/60495114
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
 ; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AN
 ; FILE REFERENCE: CL001480
 ; CURRENT APPLICATION NUMBER: US/60/495,114
 ; CURRENT FILING DATE: 2003-08-15
 ; NUMBER OF SEQ ID NOS: 91238
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2042
 ; LENGTH: 209
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-60-495-114-2042

Query Match
 Best Local Similarity 100.0%; Score 7; DB 7; Length 209;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 DAGRYLC 325
 |||||
 Db 143 DAGRYLC 149

RESULT 11
 PCT-US03-25276-32
 ; Sequence 32, Application PC/TUS0325276
 ; GENERAL INFORMATION:
 ; APPLICANT: Norris, Susan R
 ; APPLICANT: Lincoln, Kim
 ; APPLICANT: Abad, Mark Scott
 ; APPLICANT: Eilers, Robert
 ; APPLICANT: Hartsuyker, Karen Kindle
 ; APPLICANT: Hirschberg, Joseph
 ; APPLICANT: Karunanandaa, Balasubramini
 ; APPLICANT: Moshiri, Farhad
 ; APPLICANT: Stein, Joshua C.
 ; APPLICANT: Valentin, Henry E.
 ; APPLICANT: Venkatesh, Tyamagondlu V.
 ; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
 ; FILE REFERENCE: Ren-01-125
 ; CURRENT APPLICATION NUMBER: PCT/US03/25276
 ; CURRENT FILING DATE: 2003-08-05
 ; PRIOR APPLICATION NUMBER: US 60/400,689
 ; PRIOR FILING DATE: 2002-08-05
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 32
 ; LENGTH: 216
 ; TYPE: PRF
 ; ORGANISM: Synechococcus sp. -WH-8102
 PCT-US03-25276-32

Query Match
 Best Local Similarity 100.0%; Score 7; DB 1; Length 216;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 639 ALVSTLV 645

Db 157 ALVSTLV 163

RESULT 12

PCT-US02-18638A-176
; Sequence 176, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 176
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-176

Query Match 0.9%; Score 7; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632
Db 23 DSGSSEE 29

RESULT 13

PCT-US02-18638A-174
; Sequence 174, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-174

Query Match 0.9%; Score 7; DB 1; Length 300;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632
Db 23 DSGSSEE 29

RESULT 14

PCT-US03-10783-2
; Sequence 2, Application PC/TUS0310783
; GENERAL INFORMATION:
; APPLICANT: Wang, Xin Wei
; APPLICANT: Ye, Qing-Hai
; APPLICANT: Kim, Jin Woo
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; TITLE OF INVENTION: Methods of Diagnosing Potential for Metastasis or
; TITLE OF INVENTION: Developing Hepatocellular Carcinoma and of Identifying
; FILE REFERENCE: 015280-463100PC
; CURRENT APPLICATION NUMBER: PCT/US03/10783
; CURRENT FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: US 60/370,895
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: osteopontin (OPN)
PCT-US03-10783-2

Query Match 0.9%; Score 7; DB 1; Length 300;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632
Db 23 DSGSSEE 29

RESULT 15

US-60-487-610-2568
; Sequence 2568, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2568
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2568

Query Match 0.9%; Score 7; DB 7; Length 300;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632
Db 23 DSGSSEE 29

Search completed: September 9, 2003, 23:28:24
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:14:22; Search time 45 Seconds

(without alignments)
1632.731 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764

Sequence: 1 MLEVLCTCLAVFAIPATSKS.....DFLLQSTVAEAADQDQEA 764

Scoring table: OLIGO

Searched: 283308 seqs, 96168682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

1: PIR_76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	1	ORHUGS
2	21	2.7	757	1	S48841
3	21	2.7	757	1	I45956
4	17	2.2	769	1	ORRTGS
5	13	1.7	773	1	ORRBG
6	10	1.3	308	2	B84311
7	8	1.0	48	2	T35253
8	8	1.0	56	2	E64402
9	8	1.0	197	2	F71248
10	8	1.0	251	2	E90206
11	8	1.0	287	2	S55662
12	8	1.0	287	2	A82318
13	8	1.0	308	2	T45133
14	8	1.0	400	2	D64462
15	8	1.0	528	2	T21834
16	8	1.0	556	2	E75049
17	8	1.0	694	2	A69768
18	8	1.0	799	2	S18209
19	8	1.0	802	1	TVHUF4
20	8	1.0	1298	2	A48999
21	8	1.0	1363	2	I58375
22	8	1.0	1715	2	C40228
23	7	0.9	36	2	A40723
24	7	0.9	87	2	S43906
25	7	0.9	89	2	A84301
26	7	0.9	93	2	A82212
27	7	0.9	94	2	S48421
28	7	0.9	103	2	A32167
29	7	0.9	110	2	T31045

30	7	0.9	111	2	G85630	hypothetical prote
31	7	0.9	113	2	F84671	hypothetical prote
32	7	0.9	116	2	D90007	conserved hypothet
33	7	0.9	118	2	S59930	hypothetical prote
34	7	0.9	119	2	T34624	NADH2 dehydrogenas
35	7	0.9	120	2	G70700	hypothetical prote
36	7	0.9	125	2	T05445	hypothetical prote
37	7	0.9	126	2	D95410	hypothetical prote
38	7	0.9	127	2	E86158	60S ribosomal prot
39	7	0.9	128	2	H70435	hypothetical prote
40	7	0.9	128	2	H97140	probable membrane
41	7	0.9	135	2	F87515	hypothetical prote
42	7	0.9	138	2	A81264	probable periplasm
43	7	0.9	142	2	A59098	hypothetical prote
44	7	0.9	143	1	KEMS	epsilon-casein pre
45	7	0.9	146	2	E89842	hypothetical prote

ALIGNMENTS

RESULT 1

secretory component precursor [validated] - human

N:Alternate names: poly-Ig receptor; polymyric immunoglobulin receptor

N:Contains: free secretory component; transmembrane secretory component

C:Species: Homo sapiens (man)

C>Date: 28-Aug-1985 #sequence revision 23-Aug-1996 #text change 08-Dec-2000

C:Accession: A46537; A55284; I38115; A32263; S38978; S13453; A02112

R:Krajci, P.; Kvale, D.; Tasken, K.; Brandtzaeg, P.

Eur. J. Immunol. 22, 2309-2315, 1992

A:Title: Molecular cloning and exon-intron mapping of the gene encoding human tra

A:Reference number: A46537; MUID:92387236; PMID:1355431

A:Accession: A46537

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-764 <KRA>

A:Cross-references: GB:S43449; NID:9255097; PIDN:AA823176.1; PID:9255098

A:Experimental source: leukocytes

A>Note: sequence extracted from NCBI backbone (NCBI:113253)

R:Krajci, P.; Grzeschik, K.H.; Geurts van Kessel, A.H.; Olausen, B.; Brandtzaeg, J

Hum. Genet. 87, 642-648, 1991

A:Title: The human transmembrane secretory component (poly-Ig receptor): molecula

A:Reference number: A55284; MUID:92039621; PMID:1682231

A:Accession: A55284

A:Molecule type: mRNA

A:Residues: 1-764 <KRA>

A:Cross-references: GB:S62403; NID:9238235; PIDN:AA820203.1; PID:9238236

A:Experimental source: colonic adenocarcinoma cell line

A>Note: sequence extracted from NCBI backbone (NCBI:62403; NCBI:62408)

R:Piskurich, J.F.; France, J.A.; Tamer, C.M.; Willmer, C.A.; Kaetzl, C.S.; Kaetzl

Mol. Immunol. 30, 413-421, 1993

A:Title: Interferon-gamma induces polymeric immunoglobulin receptor mRNA in human

A:Reference number: I38115; MUID:93205018; PMID:8455639

A:Accession: I38115

A:Molecule type: mRNA

A:Residues: 1-764 <RES>

A:Cross-references: EMBL:X73079; NID:9456345; PIDN:CAA51532.1; PID:9456346

A>Note: submitted to the EMBL/GenBank/DBD databases by J.F. Piskurich, February

R:Krajci, P.; Solberg, R.; Sandberg, M.; Oyen, O.; Johnsen, T.; Brandtzaeg, P.

Biochem. Biophys. Res. Commun. 158, 783-789, 1989

A:Title: Molecular cloning of the human transmembrane secretory component (poly-Ig

A:Accession: A32263

A:Reference number: A32263; MUID:89149795; PMID:2920039

A:Molecule type: mRNA

A:Residues: 72-764 <KRA>

A:Cross-references: GB:M24559; NID:9514365; PIDN:AA36102.1; PID:9514366

R:Fallgren-Gebauer, E.; Gebauer, W.; Bastian, A.; Kratzin, H.D.; Elffert, H.; Zhi

Biol. Chem. Hoppe-Seyler 374, 1023-1028, 1993

A:Title: The covalent linkage of secretory component to IgA. Structure of sIgA.

A:Reference number: S38978; MUID:94121784; PMID:8292260

A:Accession: S38978

A:Molecule type: protein

A:Residues: 478-488:517-526:543-545 <PAL>
A:Note: disulfide bonds for unbound and Iga-bound forms
R:Elfert, H.; Quentlin, E.; Wiedehof, M.; Hillemeier, S.; Decker, J.; Weber, M.; Hilsch
Bioi. Chem. Hoppe-Seyler 372, 119-128, 1991
A:Title: Determination of the molecular structure of the human free secretory component.
A:Reference number: S13453; MUID:91315750; PMID:1859628
A:Accession: S13453
A:Molecule type: protein
A:Residues: 19-157, 'Q', 137-157, 'D', 159-207, 'DE', 210-228, 230-233, 'N', 235-240, 'Q', 242-261,
R:Elfert, H.; Quentlin, E.; Decker, J.; Hillemeier, S.; Hufschmidt, M.; Klingmüller, D.;
Hoppe-Seyler's Z. Physiol. Chem. 365, 1489-1495, 1984
A:Title: The primary structure of the human free secretory component and the arrangement
A:Reference number: A02112; MUID:65128981; PMID:6526384
A:Accession: A02112
A:Molecule type: protein
A:Residues: 19-157, 'D', 159-207, 'DE', 210-228, 230-233, 'N', 235-240, 'Q', 242-261,
A:Note: paper in German with English abstract
C:Comment: As a 100k transmembrane receptor for polymeric immunoglobulins, secretory com-
plicated, forms interchain disulfide bonds, undergoes proteolysis and transcytosis. Free s
C:Genetics:
A:Gene: GDB:PIGR
A:Cross-references: GDB:120290; OMIM:173880
A:Map position: 1931-1941
A:Introns: 15/1; 130/1; 349/1; 460/1; 569/1; 629/2; 670/1; 714/1; 733/3
A:Note: the first intron occurs before the initiator codon
C:Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heterod-
amers; hetero-22-mer composed of one chain of secretory component, one chain of immunog-
C:Superfamily: secretory component; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; transcyt
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-764/Product: transmembrane secretory component #status predicted <MATM>
F:19-577/Product: free secretory component #status experimental <MATF>
F:33-112/Domain: immunoglobulin homology <IM1>
F:145-222/Domain: immunoglobulin homology <IM2>
F:250-327/Domain: immunoglobulin homology <IM3>
F:364-443/Domain: immunoglobulin homology <IM4>
F:475-546/Domain: immunoglobulin homology <IM5>
F:639-661/Domain: transmembrane #status predicted <TM>
F:662-764/Domain: intracellular #status predicted <IN>
F:46-110, 56-64, 152-220, 257-323, 271-279, 311-441, 385-395, 482-544, 496-503/Disulfide bonds:
F:83, 90, 133, 186, 421, 469, 499/Binding site: carbohydrate (asn) (covalent) #status experime
F:486-520/Disulfide bonds: (in Ig-unbound form) #status experimental
F:486-520/Disulfide bonds: interchain (to Iga alpha-1 chain-192) #status experimental
F:520/Binding site: cysteine (Cys) (covalent) (in Ig-bound form) #status experimental
F:577-578/Cleavage site: Lys-Ala (unidentified proteinase) #status experimental
F:673/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 764; DB 1; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFVLTCLAVEPAISTKSPFGPEEVSNEGSNSVITCYPTSVNRRHTRKYWCROGA 60
Db 1 MLFVLTCLAVEPAISTKSPFGPEEVSNEGSNSVITCYPTSVNRRHTRKYWCROGA 60
QY 61 RGGCITLISSEGVSSKRYARALNTNPENGTFVYVIAQISODDSCKRYKGLGINSGLS 120
Db 61 RGGCITLISSEGVSSKRYARALNTNPENGTFVYVIAQISODDSCKRYKGLGINSGLS 120
QY 121 FVDSLEVSQGPLNDPKVTVYDGRVTINCPFKTEKNAKRSLKYQICILYPLVYDSS 180
Db 121 FVDSLEVSQGPLNDPKVTVYDGRVTINCPFKTEKNAKRSLKYQICILYPLVYDSS 180
QY 181 GYVNPNTGRIRLDIOGTGQLFSLVYINQLRLSDAGQYLQAGDDSNRNKNDLOYLKP 240
Db 181 GYVNPNTGRIRLDIOGTGQLFSLVYINQLRLSDAGQYLQAGDDSNRNKNDLOYLKP 240
QY 241 EPELYVEDLRGSYTFHCLGPEVANAARFLCROSSSGCNCVYVNTLGKRAPAEGRILLN 300
Db 241 EPELYVEDLRGSYTFHCLGPEVANAARFLCROSSSGCNCVYVNTLGKRAPAEGRILLN 300
QY 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIDAMOLFVNEESTIPRSPTVVK 360
Db 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIDAMOLFVNEESTIPRSPTVVK 360

Db 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLOEGSPIDAMOLFVNEESTIPRSPTVVK 360
QY 361 GVAGSSVAALCPYRNRESKSIKYWCMEGAQNGRCPLVDSQWKAQYEGRLSLEBPG 420
Db 361 GVAGSSVAALCPYRNRESKSIKYWCMEGAQNGRCPLVDSQWKAQYEGRLSLEBPG 420
QY 421 NGFTFYILNDLSDRGAFWCLNTGTLRTTYEIKIIGEPRLKYPGAVTAVLGETLVY 480
Db 421 NGFTFYILNDLSDRGAFWCLNTGTLRTTYEIKIIGEPRLKYPGAVTAVLGETLVY 480
QY 481 PCGFPEKFSSEYKWKMNNTGQALPDSDEGSKAFVNCDENSRLVSLTLNLYTRADSG 540
Db 481 PCGFPEKFSSEYKWKMNNTGQALPDSDEGSKAFVNCDENSRLVSLTLNLYTRADSG 540
QY 541 WYMGVKGQHFYGETAAYVAAVEERRAAGSRDYSKADAAPEKYLDSGFREIENKA10 600
Db 541 WYMGVKGQHFYGETAAYVAAVEERRAAGSRDYSKADAAPEKYLDSGFREIENKA10 600
QY 601 DPLFLPAEERAVATPDQADGSRASVSGSSEEGGSSRLVSTLVPLGLVLAAGVAVGY 660
Db 601 DPLFLPAEERAVATPDQADGSRASVSGSSEEGGSSRLVSTLVPLGLVLAAGVAVGY 660
QY 661 ARAHRRKNDVRSIRSYRTDISMSDENSREFCANMNGASSITQETSLGKBEFVATTE 720
Db 661 ARAHRRKNDVRSIRSYRTDISMSDENSREFCANMNGASSITQETSLGKBEFVATTE 720
QY 721 STTEKEPKKAKSSKEEEMAYKDFLLQSTVAAEAODGPOEA 764
Db 721 STTEKEPKKAKSSKEEEMAYKDFLLQSTVAAEAODGPOEA 764

RESULT 2
S48841
secretory component precursor - bovine
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
C:Contains: free secretory component; transmembrane secretory component
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 10-Sep-1999
R:Vermeer, H.; Warmerdam, G.W.; de Boer, H.A.; Verbeet, M.P.H.
submitted to the EMBL Data Library, September 1994
A:Description: The cloning, tissue specific expression and interspecies sequence c
A:Reference number: S48841
A:Accession: S48841
A:Molecule type: mRNA
A:Residues: 1-757 <VER>
A:Cross-references: EMBL:X81371; NID:9563340; PIDN:CAA57136.1; PID:9563341
C:Superfamily: secretory component; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; tr-
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-757/Product: transmembrane secretory component #status predicted <MATM>
F:33-112/Domain: immunoglobulin homology <IM1>
F:145-222/Domain: immunoglobulin homology <IM2>
F:250-328/Domain: immunoglobulin homology <IM3>
F:365-444/Domain: immunoglobulin homology <IM4>
F:476-547/Domain: immunoglobulin homology <IM5>
F:631-653/Domain: transmembrane #status predicted <TM>
F:654-757/Domain: intracellular #status predicted <IN>
F:46-110, 56-64, 152-220, 166-173, 257-324, 271-279, 370-440, 384-394, 481-543, 495-502/Dis-
F:83, 420, 468/Binding site: carbohydrate (asn) (covalent) #status predicted
F:665/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.7%; Score 21; DB 1; Length 757;
Best Local Similarity 100.0%; Pred. No. 2,3e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 CYPPTSVNRRHTRKYWCROGA 60
Db 40 CYPPTSVNRRHTRKYWCROGA 60

RESULT 3

145956
 polymeric immunoglobulin receptor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 23-Jul-1999
 C:Accession: 145956
 R:Kuslseth, M.A.; Krajaei, P.; Myklebost, O.; Rognes, S.
 DNA Cell Biol. 14, 251-256, 1995
 A:Title: Cloning and characterization of two forms of bovine polymeric immunoglobulin re
 A:Reference number: 145956; MUID:95186063; PMID:7880445
 A:Accession: 145956
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-757 <KUL>
 A:Cross-references: GB:L04797; NID:9388279; PIDN:AA041620.1; PID:9388280
 C:Superfamily: secretory component; immunoglobulin homology
 F:145-222/Domain: immunoglobulin homology <IM4>

Query Match 2.7%; Score 21; DB 1; Length 757;
 Best Local Similarity 100.0%; Pred. No. 2,38-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 CYPPTSVNRHRRKRWCRGA 60
 |||||||
 Db 40 CYPPTSVNRHRRKRWCRGA 60

RESULT 4

ORRG

secretory component precursor - rat
 N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
 C:Contains: free secretory component; transmembrane secretory component
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-Sep-1990 #sequence_revision 23-Aug-1996 #text_change 16-Jul-1999
 C:Accession: S05407; S54731
 R:Banfill, G.; Brake, B.; Braghetta, P.; Luzzo, J.P.; Stanley, K.K.
 FEBS Lett. 254, 177-183, 1989
 A:Title: Intracellular targeting signals of polymeric immunoglobulin receptors are high
 A:Reference number: S05407; MUID:89378226; PMID:2776882
 A:Accession: S05407
 A:Molecule type: mRNA
 A:Residues: 1-769 <BAN>
 A:Cross-references: EMBL:X15741; NID:956464; PIDN:CAA33758.1; PID:956465
 C:Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heterod
 amers; hetero-22-mer composed of one chain of secretory component, one chain of immunog
 C:Superfamily: secretory component; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; transcyt
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-769/Product: transmembrane secretory component #status predicted <MATN>
 F:19-582/Product: free secretory component #status predicted <MATF>
 F:33-112/Domain: immunoglobulin homology <IM1>
 F:145-222/Domain: immunoglobulin homology <IM2>
 F:250-326/Domain: immunoglobulin homology <IM3>
 F:363-442/Domain: immunoglobulin homology <IM4>
 F:477-548/Domain: immunoglobulin homology <IM5>
 F:644-666/Domain: transmembrane #status predicted <TM>
 F:667-769/Domain: intracellular #status predicted <INT>
 F:40-110,56-64,152-220,370-440,384-546,498-505/Disulfide bonds: #status predicted
 F:90,135,471/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:488-522/Disulfide bonds: (in Ig-unbound form) #status predicted
 F:488/Disulfide bonds: Interchain (to Iga alpha-1 chain-192) #status predicted
 F:522/Binding site: cysteine (cys) (in Ig-bound form) #status predicted
 F:678/Binding site: phosphate (ser) (covalent) #status predicted

Query Match 2.2%; Score 17; DB 1; Length 769;
 Best Local Similarity 100.0%; Pred. No. 2,66-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 STLVPLGLVAVAV 658
 |||||||
 Db 647 STLVPLGLVAVAV 663

RESULT 5

ORRG

secretory component precursor - rabbit
 N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
 C:Contains: free secretory component; transmembrane secretory component
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
 C:Accession: A02111; A28077
 R:Mostov, K.E.; Friedlander, M.; Blobel, G.
 Nature 308, 37-43, 1984
 A:Title: The receptor for trans epithelial transport of Iga and Igm contains multi
 A:Reference number: A02111; MUID:84142246; PMID:6322002
 A:Accession: A02111
 A:Molecule type: mRNA
 A:Residues: 1-773 <MOS>
 A:Cross-references: GB:X00412; GB:K01291; NID:91595; PIDN:CAA25118.1; PID:91596
 A:Note: the authors translated the codon ACC for residue 54 as Asn
 R:Frutiger, S.; Hughes, G.J.; Hanly, W.C.; Jaton, J.C.
 J. Biol. Chem. 263, 8120-8125, 1988
 A:Title: Rabbit secretory components of different allotypes vary in their carbohy
 A:Reference number: A28077; MUID:88228032; PMID:3131339
 A:Accession: A28077

A:Molecule type: protein
 A:Residues: 87-114,410-424 <FRU>
 C:Comment: This receptor binds polymeric Iga and Igm at the basolateral surface o
 process, cleavage occurs to separate the extracellular portion, also known as the
 C:Comment: The five domains exhibit homology with immunoglobulin V regions. The s
 C:Comment: Alternative splicing in the extracellular domain leads to high or low
 C:Superfamily: secretory component; immunoglobulin homology
 C:Keywords: alternative splicing; duplication; glycoprotein; immunoglobulin recep
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-773/Product: transmembrane secretory component #status predicted <MATN>
 F:19-575/Product: free secretory component #status predicted <MATF>
 F:30-647/Domain: intracellular #status predicted <EXT>
 F:39-117/Domain: immunoglobulin homology <IM1>
 F:148-227/Domain: immunoglobulin homology <IM2>
 F:253-326/Domain: immunoglobulin homology <IM3>
 F:362-440/Domain: immunoglobulin homology <IM4>
 F:471-540/Domain: immunoglobulin homology <IM5>
 F:648-670/Domain: transmembrane #status predicted <TM>
 F:671-773/Domain: intracellular #status predicted <INT>
 F:46-115,155-225,260-324,369-438,478-538/Disulfide bonds: #status predicted
 F:108/Binding site: carbohydrate (asn) (covalent) (partial) #status experimental
 F:418/Binding site: carbohydrate (asn) (covalent) #status experimental

Query Match 1.7%; Score 13; DB 1; Length 773;
 Best Local Similarity 100.0%; Pred. No. 0,00029;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 EPRKAKRSSKEA 739
 |||||||
 Db 736 EPRKAKRSSKEA 748

RESULT 6

B84311

hypothetical protein Vng1578h [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: B84311
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lu
 ; Leitthausser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.
 Jung, K.H.; Alam, M.; Freilich, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T
 A:Title: Genome sequence of Halobacterium species NRC-1
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: B84311
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-308 <STO>
 A:Cross-references: GB:AE004437; NID:q10581062; PIDN:AA019854.1; GSPDB:GN00138
 C:Genetics:

A:Gene: VNG1578H
C:Superfamily: cobalamin biosynthesis protein D

Query Match 1.38; Score 10; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 652 AVGAVANGVA 661
DB 5 AVGAVANGVA 14

RESULT 7
T35253
Small hypothetical protein SC5F2A.11 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35253
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A:Reference number: 221573
A:Accession: T35253
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-48 <OLIT>
A:Cross-references: EMBL:AL049587; PIDN:CAB40678.1; GSPDB:GN00070; SCOEDB:SC5F2A.11
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC5F2A.11

Query Match 1.0%; Score 8; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 648 GLVLAVGA 655
DB 24 GLVLAVGA 31

RESULT 8
E64402
hypothetical protein Mj0821 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Oct-1999
C:Accession: E64402
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: E64402
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-56 <BULT>
A:Cross-references: GB:U67526; GB:L77117; NID:q1591508; PIDN:AB98833.1; PID:q1499645; T
C:Genetics:
A:Map position: REV742767-742597

Query Match 1.0%; Score 8; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 457 IIEGPNL 464
DB 16 IIEGPNL 23

RESULT 9
F71248
probable proteasome beta subunit - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: F71248
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushiida, N.; DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermoph
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: F71248
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-197 <KAW>
A:Cross-references: GB:AP000001; NID:q3236128; PIDN:BAA29317.1; PID:q3256634
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Ge
C:Genetics:
A:Gene: PH0245
C:Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 1.0%; Score 8; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 566 KAAGSRDV 573
DB 166 KAAGSRDV 173

RESULT 10
E90206
hypothetical protein hisf [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: E90206
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Jorg, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Red
arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90206
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-251 <KUR>
A:Cross-references: GB:AE006641; NID:q13813762; PIDN:AAK40908.1; GSPDB:GN00155
C:Genetics:
A:Gene: hisf
C:Superfamily: cyclase hisf

Query Match 1.0%; Score 8; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 574 SLAKADAA 581
DB 215 SLAKADAA 222

RESULT 11
S55662
legument protein 67 - equine herpesvirus 2
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S55662
R:Telford, E.A.R.; Watson, M.S.; Auld, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501; PMID:7783207
A:Accession: S55662
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-287 <TEL>
A:Cross-references: GB:U20824; NID:q695172; PIDN:AAK13855.1; PID:q695240
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

C:Superfamily: human herpesvirus 4 BRF1 protein

Query Match 1.0%; Score 8; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 VLAVGAVA 657
Db 268 VLAVGAVA 275

RESULT 12

conserved hypothetical protein VC0480 [Imported] - Vibrio cholerae (strain N16961 serog
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82318

R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: A82318
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-287 <HEI>
A:Cross-references: GB:AE004134; GB:AE003852; NID:9654900; PIDN:APF93653.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0480
A:Map position: 1
C:Superfamily: Escherichia coli hypothetical 30.9K protein (sbm-fba intergenic region)

Query Match 1.0%; Score 8; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 VVKGVAGS 365
Db 50 VVKGVAGS 57

RESULT 13

T45133
restriction endonuclease Mami [Imported] - Microbacterium ammoniaphilum
N:Alternate names: restriction-modification system Mami
C:Species: Microbacterium ammoniaphilum
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 16-Feb-2001
C:Accession: T45133

R:Stiebel, H.M.; Seeber, S.; Jarsch, M.; Kessler, C.
Gene 172, 41-46, 1996

A:Title: Cloning and characterization of the Mami restriction-modification system from M
A:Reference number: Z22923; MUID:96257250; PMID:8654988

A:Accession: T45133

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-308 <STR>

A:Cross-references: EMBL:X79027; NID:9984667; PIDN:CA55648.1; PID:9984669

A:Experimental source: ATCC 15354

C:Genetics:

A:Gene: mamiR

C:Superfamily: Microbacterium ammoniaphilum restriction endonuclease Mami

Query Match 1.0%; Score 8; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 DVVNTLG 287
Db 200 DVVNTLG 207

RESULT 14

D64462
hypothetical protein MJ1301 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: D64462
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; I
Reich, C.I.; Overbeek, R.; Kirsch, E.F.; Weissbrock, R.G.; Merrick, J.M.; Glode
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Wk
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jani
A:Reference number: A64300; MUID:96337999; PMID:868087

A:Accession: D64462

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-400 <BUL>

A:Cross-references: GB:U67571; GB:U77117; NID:q1591939; PIDN:AAB99326.1; PID:q159;

C:Genetics:

A:Map position: REV1250122-1248920

C:Superfamily: ATP-binding protein PAB1945

Query Match 1.0%; Score 8; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 IIEGEPML 464
Db 23 IIEGEPML 30

RESULT 15

T21834
hypothetical protein F36A2.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21834

R:Lennard, N.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19476

A:Accession: T21834

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-528 <WIL>

A:Cross-references: EMBL:Z81077; PIDN:CAB03071.1; GSPDB:GN00019; CESP:F36A2.4

A:Experimental source: Clone F36A2

C:Genetics:

A:Gene: CESP:F36A2.4

A:Map position: 1

Query Match 1.0%; Score 8; DB 2; Length 528;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 VANVAKFL 270
Db 203 VANVAKFL 210

Search completed: September 9, 2003, 23:18:39
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:07:32 : Search time 26 Seconds

(Without alignments)
1381.862 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764
Sequence: 1 MLFVLTCLLAVPAISTKS.....DFLLQSTVAEAQDQGEA 764

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	1	PIGR_HUMAN
2	757	2.7	757	1	PIGR_BOVIN
3	769	2.2	769	1	PIGR_RAT
4	16	2.1	771	1	PIGR_MOUSE
5	13	1.7	773	1	PIGR_RABIT
6	11	1.4	102	1	PIGR_PIG
7	8	1.0	56	1	Y821_METJA
8	8	1.0	251	1	HIS6_SULSO
9	8	1.0	308	1	T2M1_MICAM
10	8	1.0	400	1	YD01_METJA
11	8	1.0	556	1	SYFB_PYRAB
12	8	1.0	556	1	SYFB_PYRAB
13	8	1.0	563	1	B2NB_PSEFL
14	8	1.0	802	1	FCR4_HUMAN
15	8	1.0	808	1	FCR4_MOUSE
16	8	1.0	1298	1	VGR3_HUMAN
17	8	1.0	1363	1	VGR3_MOUSE
18	8	1.0	1712	1	NX2A_HUMAN
19	8	1.0	1715	1	NX2A_RAT
20	7	0.9	94	1	YF8_YEAST
21	7	0.9	110	1	YB1_STRAU
22	7	0.9	120	1	Y025_MYCTU
23	7	0.9	127	1	R22A_ARATH
24	7	0.9	128	1	WNT8_THUTH
25	7	0.9	128	1	YF70_AQUAE
26	7	0.9	143	1	CAS2_MOUSE
27	7	0.9	153	1	HXC6_SHEEP
28	7	0.9	163	1	IPYR_STRCO
29	7	0.9	209	1	PGPI_HUMAN
30	7	0.9	209	1	PGPI_MOUSE
31	7	0.9	228	1	MTH2_HAENA
32	7	0.9	233	1	US56_HSVIE
33	7	0.9	233	1	HXC6_MOUSE

34	7	0.9	247	1	HXA4_HETFR	091a22 heterodontu
35	7	0.9	250	1	COX2_NEUCR	P00411 neurospora
36	7	0.9	255	1	HXD4_HUMAN	P09016 homo sapien
37	7	0.9	257	1	MOTA_HELPY	P56426 helicobacte
38	7	0.9	259	1	YGFW_ECOLI	046813 escherichia
39	7	0.9	277	1	KD5A_BROBE	089h11 brucella me
40	7	0.9	277	1	KD5A_RHILU	086m26 rhizobium l
41	7	0.9	278	1	Y110_MYCPN	P75523 mycoplasma
42	7	0.9	279	1	HHPX_NEIMB	09K006 neisseria m
43	7	0.9	280	1	KD5A_RHIME	092q99 rhizobium m
44	7	0.9	282	1	KD5A_AGRF5	08ufh3 agrobacteri
45	7	0.9	285	1	SYGB_HAENA	030836 haemophilus

ALIGNMENTS

RESULT 1	PIGR_HUMAN	STANDARD:	PRT:	764 AA.
AC	PIGR_HUMAN			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)			
DE	[Contains: Secretory component].			
GN	PIGR			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92039621; PubMed=1682231;			
RA	Krajci P., Grzeschik K.H., Geurts van Kessel A.H., Olaisen B.,			
RA	Brandtzaeg P.;			
RT	"The human transmembrane secretory component (poly-Ig receptor):			
RT	molecular cloning, restriction fragment length polymorphism and			
RT	chromosomal sublocalization.";			
RL	Hum. Genet. 87:642-648(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92387236; PubMed=1355431;			
RA	Krajci P., Kvale D., Tasken K., Brandtzaeg P.;			
RT	"Molecular cloning and exon-intron mapping of the gene encoding human			
RT	transmembrane secretory component (the poly-Ig receptor).";			
RL	Eur. J. Immunol. 22:2309-2315(1992).			
RN	[3]			
RP	SEQUENCE OF 72-764 FROM N.A.			
RX	MEDLINE=89149795; PubMed=2920039;			
RA	Krajci P., Solberg R., Sandberg M., Oyen O., Jahnson T.,			
RA	Brandtzaeg P.;			
RT	"Molecular cloning of the human transmembrane secretory component			
RT	(poly-Ig receptor) and its mRNA expression in human tissues.";			
RL	Biochem. Biophys. Res. Commun. 156:783-789(1989).			
RN	[4]			
RP	SEQUENCE OF 19-577, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.			
RX	MEDLINE=85128981; PubMed=6526384;			
RA	Eiffert H., Quentin E., Decker J., Hillemeir S., Hufschmidt M.,			
RA	Klingmüller D., Weber M.H., Hilschmann N.;			
RT	"The primary structure of human free secretory component and the			
RT	arrangement of disulfide bonds.";			
RL	Hoppe-Seyler's Z. Physiol. Chem. 365:1489-1495(1984).			
RN	[5]			
RP	SEQUENCE OF 19-577.			
RX	MEDLINE=91315750; PubMed=1859628;			
RA	Eiffert H., Quentin E., Wiederhold M., Hillemeir S., Decker J.,			
RA	Weber M., Hilschmann N.;			
RT	"Determination of the molecular structure of the human free secretory			
RT	component.";			
RL	Biol. Chem. Hoppe-Seyler 372:119-126(1991).			
RN	[6]			
RP	SEQUENCE OF 118-138; 212-230; 232-268; 273-288 AND 578-603.			

RX MEDLINE-92379357; PubMed-9237679;
 RA Hughes G.J., Futcher S., Savoy L.-A., Reason A.J., Morris H.R.,
 RA Jaton J.-C.;
 RT "Human free secretory component is composed of the first 585 amino
 RT acid residues of the polymeric immunoglobulin receptor.";
 RL FEBS Lett. 410:443-446(1997).
 CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
 CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
 CC TRANSMEMBRANE SEGMENT.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
 CC -!- SIMILARITY: Contains 5 Immunoglobulin-like V-type domains.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: S62403; AAB20203.1; -;
 DR EMBL: S43443; AAB23176.1; -;
 DR EMBL: S43437; AAB23176.1; JOINED.
 DR EMBL: S43441; AAB23176.1; JOINED.
 DR EMBL: S43442; AAB23176.1; JOINED.
 DR EMBL: S43443; AAB23176.1; JOINED.
 DR EMBL: S43444; AAB23176.1; JOINED.
 DR EMBL: S43445; AAB23176.1; JOINED.
 DR EMBL: S43446; AAB23176.1; JOINED.
 DR EMBL: S43447; AAB23176.1; JOINED.
 DR EMBL: S43448; AAB23176.1; JOINED.
 DR EMBL: M24559; AAA36102.1; -;
 DR EMBL: A52091; CAA03384.1; -;
 DR PIR: A46537; QRMUS3;
 DR GlycosultedB; P01833; -;
 DR Genew; HGNC:8968; PIGR.
 DR MIM: 173880; -;
 DR GO: GO:0005887; C: Integral to plasma membrane; TAS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig; 5.
 DR SMART: SM00409; IG; 5.
 DR PROSITE: PS50835; IG LIKE; 2.
 KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
 KW Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 764 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
 FT CHAIN 19 603 SECRETORY COMPONENT.
 FT DOMAIN 19 638 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 639 661 POTENTIAL.
 FT DOMAIN 662 764 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 19 120 IG-LIKE V-TYPE 1.
 FT DOMAIN 145 237 IG-LIKE V-TYPE 2.
 FT DOMAIN 250 352 IG-LIKE V-TYPE 3.
 FT DOMAIN 364 458 IG-LIKE V-TYPE 4.
 FT DOMAIN 462 561 IG-LIKE V-TYPE 5.
 FT DISULFID 40 110
 FT DISULFID 56 64
 FT DISULFID 152 220
 FT DISULFID 257 325
 FT DISULFID 271 279
 FT DISULFID 371 441
 FT DISULFID 385 395
 FT DISULFID 482 544
 FT DISULFID 486 520
 FT DISULFID 496 503
 FT CARBOHYD 83 83
 FT CARBOHYD 90 90
 FT CARBOHYD 135 135
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)

FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .)
 FT VARIANT 580 580 A -> V.
 FT VARIANT 580 580 /PRTD-VAR. 003920.
 FT CONFLICT 136 136 D -> Q (IN REF. 4 AND 5).
 FT CONFLICT 158 158 N -> D (IN REF. 4 AND 5).
 FT CONFLICT 208 209 NO -> DE (IN REF. 4 AND 5).
 FT CONFLICT 229 229 MISSING (IN REF. 4 AND 5).
 FT CONFLICT 234 234 D -> N (IN REF. 4 AND 5).
 FT CONFLICT 241 241 E -> Q (IN REF. 4 AND 5).
 FT CONFLICT 262 262 E -> Q (IN REF. 4 AND 5).
 FT CONFLICT 280 280 D -> N (IN REF. 4 AND 5).
 FT CONFLICT 392 392 N -> D (IN REF. 4 AND 5).
 FT CONFLICT 500 500 N -> D (IN REF. 4 AND 5).
 SO SEQUENCE 764 AA; 83313 MW; 916B3E662C339950 CRC64;
 Query Match 100.0%; Score 764; DB 1; Length 764;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLFVLTCLAVPAISTKSPIFGPEEVNSVSGNSVSYNCYPTSVNRRTRKYMGROGA 60
 DB 1 MLFVLTCLAVPAISTKSPIFGPEEVNSVSGNSVSYNCYPTSVNRRTRKYMGROGA 60
 QY 61 RGGCITLISSEGYSSKYAGRANLTFNPENGTFFVNIADLSQDSGRYKCGLGINSRGLS 120
 DB 61 RGGCITLISSEGYSSKYAGRANLTFNPENGTFFVNIADLSQDSGRYKCGLGINSRGLS 120
 QY 121 FDIYSLESGPGGLINTKTYTVDLGRVYINCFFKRNMAKRSYKQJGLYVLYIDS 180
 DB 121 FDIYSLESGPGGLINTKTYTVDLGRVYINCFFKRNMAKRSYKQJGLYVLYIDS 180
 QY 121 FDIYSLESGPGGLINTKTYTVDLGRVYINCFFKRNMAKRSYKQJGLYVLYIDS 180
 DB 121 FDIYSLESGPGGLINTKTYTVDLGRVYINCFFKRNMAKRSYKQJGLYVLYIDS 180
 QY 181 GYVNPVYTGRIKRIIDIGTGGLFSVYINOLRLSDAGQYLCQAGDSDSNKKNADLYLKP 240
 DB 181 GYVNPVYTGRIKRIIDIGTGGLFSVYINOLRLSDAGQYLCQAGDSDSNKKNADLYLKP 240
 QY 181 GYVNPVYTGRIKRIIDIGTGGLFSVYINOLRLSDAGQYLCQAGDSDSNKKNADLYLKP 240
 DB 181 GYVNPVYTGRIKRIIDIGTGGLFSVYINOLRLSDAGQYLCQAGDSDSNKKNADLYLKP 240
 QY 241 EPELVYEDLRSVTFICALGPEVANYAKFLCROSSSENCDDVYNTLGKRAPAFEGRIILN 300
 DB 241 EPELVYEDLRSVTFICALGPEVANYAKFLCROSSSENCDDVYNTLGKRAPAFEGRIILN 300
 QY 241 EPELVYEDLRSVTFICALGPEVANYAKFLCROSSSENCDDVYNTLGKRAPAFEGRIILN 300
 DB 241 EPELVYEDLRSVTFICALGPEVANYAKFLCROSSSENCDDVYNTLGKRAPAFEGRIILN 300
 QY 301 POKKDSFSVYITGLKREDAGRYLCGAHSGQLQDESPLOAMOLFVNEESTIPRSPYVK 360
 DB 301 POKKDSFSVYITGLKREDAGRYLCGAHSGQLQDESPLOAMOLFVNEESTIPRSPYVK 360
 QY 301 POKKDSFSVYITGLKREDAGRYLCGAHSGQLQDESPLOAMOLFVNEESTIPRSPYVK 360
 DB 301 POKKDSFSVYITGLKREDAGRYLCGAHSGQLQDESPLOAMOLFVNEESTIPRSPYVK 360
 QY 361 GVAGSSVAVLCPRNKRRESKSIKYKWLMEGQNGRCPLVDSBGMVKAQYGRSLSEEPG 420
 DB 361 GVAGSSVAVLCPRNKRRESKSIKYKWLMEGQNGRCPLVDSBGMVKAQYGRSLSEEPG 420
 QY 421 NGFTVYILNQLTSRDAGFYWCLTNGDTLMRTVEIKIIEGPNLKYPGNVTAVLGELKY 480
 DB 421 NGFTVYILNQLTSRDAGFYWCLTNGDTLMRTVEIKIIEGPNLKYPGNVTAVLGELKY 480
 QY 481 PCGFPECKFSSEYKWKCKNNNTGQALPQSDGEGSKAFVNCNDENSRVLSLTNLVTRADEG 540
 DB 481 PCGFPECKFSSEYKWKCKNNNTGQALPQSDGEGSKAFVNCNDENSRVLSLTNLVTRADEG 540
 QY 481 PCGFPECKFSSEYKWKCKNNNTGQALPQSDGEGSKAFVNCNDENSRVLSLTNLVTRADEG 540
 DB 481 PCGFPECKFSSEYKWKCKNNNTGQALPQSDGEGSKAFVNCNDENSRVLSLTNLVTRADEG 540
 QY 541 WYMGYKQGFYGETAAVYAVEEERKAAGSRDYSLAKADAPAEKYLDGSGFREIENKAIO 600
 DB 541 WYMGYKQGFYGETAAVYAVEEERKAAGSRDYSLAKADAPAEKYLDGSGFREIENKAIO 600
 QY 541 WYMGYKQGFYGETAAVYAVEEERKAAGSRDYSLAKADAPAEKYLDGSGFREIENKAIO 600
 DB 541 WYMGYKQGFYGETAAVYAVEEERKAAGSRDYSLAKADAPAEKYLDGSGFREIENKAIO 600
 QY 601 DPLRFAEERKAVADTRDQADGSRASVDSGSEEGGSSRALVSTLVPLGLVLAAGAAYGV 660
 DB 601 DPLRFAEERKAVADTRDQADGSRASVDSGSEEGGSSRALVSTLVPLGLVLAAGAAYGV 660
 QY 601 DPLRFAEERKAVADTRDQADGSRASVDSGSEEGGSSRALVSTLVPLGLVLAAGAAYGV 660
 DB 601 DPLRFAEERKAVADTRDQADGSRASVDSGSEEGGSSRALVSTLVPLGLVLAAGAAYGV 660
 QY 661 ARAHRKNDVRSVIRSTRDISDPSDENREFGANDNMGASSTIOETSLGCKEFAATTE 720
 DB 661 ARAHRKNDVRSVIRSTRDISDPSDENREFGANDNMGASSTIOETSLGCKEFAATTE 720
 QY 721 STTEKPEPKAKRSSKEEAEAMAYKDFLLOSTYAAEADQDPOEA 764
 DB 721 STTEKPEPKAKRSSKEEAEAMAYKDFLLOSTYAAEADQDPOEA 764
 QY 721 STTEKPEPKAKRSSKEEAEAMAYKDFLLOSTYAAEADQDPOEA 764
 DB 721 STTEKPEPKAKRSSKEEAEAMAYKDFLLOSTYAAEADQDPOEA 764

FT	CHAIN	19	757	POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT	CHAIN	19	599	SECRETORY COMPONENT (BY SIMILARITY).
FT	DOMAIN	19	632	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	633	653	POTENTIAL.
FT	DOMAIN	654	757	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	19	126	IG-LIKE V-TYPE 1.
FT	DOMAIN	145	237	IG-LIKE V-TYPE 2.
FT	DOMAIN	250	341	IG-LIKE V-TYPE 3.
FT	DOMAIN	353	457	IG-LIKE V-TYPE 4.
FT	DOMAIN	461	560	IG-LIKE V-TYPE 5.
FT	DISULFID	40	110	BY SIMILARITY.
FT	DISULFID	56	64	BY SIMILARITY.
FT	DISULFID	152	220	BY SIMILARITY.
FT	DISULFID	257	324	BY SIMILARITY.
FT	DISULFID	271	279	BY SIMILARITY.
FT	DISULFID	370	440	BY SIMILARITY.
FT	DISULFID	384	394	BY SIMILARITY.
FT	DISULFID	481	543	BY SIMILARITY.
FT	DISULFID	485	519	BY SIMILARITY.
FT	DISULFID	495	502	BY SIMILARITY.
FT	CARBOHYD	83	83	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	420	420	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	468	468	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	MOD_RES	727	727	PHOSPHORYLATION (BY SIMILARITY).
FT	VARSPIC	129	346	Missing (In Isoform Short).
FT	VARSPIC	129	346	/Ftd-VSP-002547.
FT	VARIANT	29	29	T -> S.
FT	VARIANT	142	142	V -> I.
FT	VARIANT	404	404	I -> M.
FT	VARIANT	413	413	A -> V.
FT	VARIANT	435	435	T -> A.
SO	SEQUENCE	757 AA;	82434 MW;	DCED67PDD6A6E6C6 CRC64;
Query Match 2.7%; Score 21; DB 1; Length 757;				
Best Local Similarity 100.0%; Pred. No. 5.2e-13;				
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	40	CYPPTSVNRRTRKRYCROGA	60	
	40	CYPPTSVNRRTRKRYCROGA	60	
Db	40	CYPPTSVNRRTRKRYCROGA	60	
RESULT 3				
PIGR_RAT	PIGR_RAT	STANDARD;	PRT;	769 AA.
ID	PIGR_RAT	STANDARD;	PRT;	769 AA.
AC	P15083;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)			
DE	[contains: Secretory component].			
GN	PIGR.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID:10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver.			
RX	MEDLINE=8937826; PubMed=2776882;			
RA	Bailling G., Brake B., Braghetta P., Luzzo J.P., Stanley K.K.;			
RT	"Intracellular targeting signals of polymeric immunoglobulin			
RT	receptors are highly conserved between species.";			
RL	FEBS Lett. 254:177-183(1989).			
CC	-1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE			
CC	BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN			
CC	TRANSFORMED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.			
CC	DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE			
CC	EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE			
CC	TRANSMEMBRANE SEGMENT.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.			
CC	-1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X15741; CAA33758.1; -
PIR: S05407; ORRTGS.
InterPro: IPR007110; Ig-like.
InterPro: IPR003599; Ig.
InterPro: IPR003006; Ig_MHC.
Pfam: PF00047; Ig; 5.
SMART: SM00409; Ig; 5.
PROSITE: PS50835; IG-LIKE; 2.
Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
CHAIN 1 18
FT CHAIN 19 769 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 611 SECRETORY COMPONENT.
FT DOMAIN 19 643 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 644 666 POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 667 769 IG-LIKE V-TYPE 1.
FT DOMAIN 21 126 IG-LIKE V-TYPE 2.
FT DOMAIN 135 237 IG-LIKE V-TYPE 3.
FT DOMAIN 353 457 IG-LIKE V-TYPE 4.
FT DOMAIN 463 563 IG-LIKE V-TYPE 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 769 AA; 84798 MW; 5F849303400255A7 CRC64;

Query Match 2.2%; Score 17; DB 1; Length 769;
Best Local Similarity 100.0%; Pred. No. 7e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 642 STLVPLGLVAVGAVAV 658
D6 647 STLVPLGLVAVGAVAV 663

RESULT 4
PIGR_MOUSE STANDARD: PRT: 771 AA.
ID PIGR_MOUSE 070570;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PolymERIC-Immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=95138517; PubMed=7836758;
RA Pliskulich J.E., Blanchard M.H., Youngman K.R., France J.A.,
RA Kaetzel C.S.;
RT "Molecular cloning of the mouse polymeric Ig receptor. Functional
RT regions of the molecule are conserved among five mammalian species."
RL J. Immunol. 154:1735-1747(1995).
RN (2)
RN SEQUENCE FROM N.A.

RC STRAIN=129/SVJ;
RX MEDLINE=98072444; PubMed=9409786;
RA Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;
RT "Genomic cloning and structural analysis of the murine polymeric
RT receptor (PIGR) gene and promoter region."
RL Gene 201:189-197(1997).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Liver;
RA de Groot N., Vollebregt E., Lee S.H., Verbeet M.P., de Boer H.A.;
RT "Molecular cloning and exon-intron organization of the gene encoding
RT the murine polymeric immunoglobulin receptor."
RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT (BY SIMILARITY). ALSO SECRETED.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: U06431; AA67440.1; -
EMBL: U83434; AAC53585.1; JOINED.
EMBL: U83427; AAC53585.1; JOINED.
EMBL: U83428; AAC53585.1; JOINED.
EMBL: U83429; AAC53585.1; JOINED.
EMBL: U83430; AAC53585.1; JOINED.
EMBL: U83431; AAC53585.1; JOINED.
EMBL: U83432; AAC53585.1; JOINED.
EMBL: U83433; AAC53585.1; JOINED.
EMBL: Y16524; CAA76272.1; JOINED.
EMBL: Y16525; CAA76272.1; JOINED.
EMBL: Y16526; CAA76272.1; JOINED.
EMBL: Y16527; CAA76272.1; JOINED.
EMBL: Y16528; CAA76272.1; JOINED.
EMBL: Y16529; CAA76272.1; JOINED.
EMBL: Y16530; CAA76272.1; JOINED.
EMBL: Y16531; CAA76272.1; JOINED.
EMBL: Y16532; CAA76272.1; JOINED.
MGD: MGI:103080; PIGR.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00409; Ig; 3.
DR PROSITE: PS50835; IG-LIKE; 3.
Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
CHAIN 1 18
FT CHAIN 19 771 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 611 SECRETORY COMPONENT.
FT DOMAIN 19 645 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 646 668 POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 669 771 IG-LIKE V-TYPE 1.
FT DOMAIN 21 120 IG-LIKE V-TYPE 2.
FT DOMAIN 135 237 IG-LIKE V-TYPE 3.
FT DOMAIN 353 457 IG-LIKE V-TYPE 4.
FT DOMAIN 463 563 IG-LIKE V-TYPE 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT DISULFID 484 546 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 147 147 N-LINKED (GLCNAAC. . .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAAC. . .) (POTENTIAL).
FT CONFLICT 159 159 A -> V (IN REF. 1).
FT CONFLICT 396 396 V -> A (IN REF. 1).
FT CONFLICT 620 620 G -> R (IN REF. 1).
SQ SEQUENCE 771 AA; 84598 MW; 78C81302EC710730 CRC64;

Query Match 2.1%; Score 16; DB 1; Length 771;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 45 TSVNRHRRKWCROGA 60
    |||||
Db 45 TSVNRHRRKWCROGA 60

RESULT 5
PIGR_RABIT STANDARD; PRT; 773 AA.
AC P01812;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=84142246; PubMed=6322002;
RA Mostov K.E., Friedlander M., Blodel G.;
RT "The receptor for transepithelial transport of IgA and Igm contains
RT multiple immunoglobulin-like domains.";
RL Nature 308:37-43(1984).
RN [2]
RP SEQUENCE OF 87-114 AND 410-428.
RX MEDLINE=8628032; PubMed=313139;
RA Frutiger S., Hughes G.J., Hanly W.C., Jaton J.-C.;
RT "Rabbit secretory components of different allotypes vary in their
RT carbohydrate content and their sites of N-linked glycosylation.";
RL J. Biol. Chem. 263:8120-8125(1988).
RN [3]
RP FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
RN BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
RN TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
RN DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
RN EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
RN TRANSMEMBRANE SEGMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -1- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF ALLOTYPIC T62.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00412; CAA25118.1; .
DR PIR; A02111; QRRBG.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 5.
DR PROSITE; PS50835; IG-LIKE; 3.

```

```

KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
KW Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 773 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 615 SECRETORY COMPONENT.
FT DOMAIN 19 647 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 131 IG-LIKE V-TYPE 1.
FT DOMAIN 138 232 IG-LIKE V-TYPE 2.
FT DOMAIN 333 340 IG-LIKE V-TYPE 3.
FT DOMAIN 352 455 IG-LIKE V-TYPE 4.
FT DOMAIN 461 557 IG-LIKE V-TYPE 5.
FT DISULFID 155 225 POTENTIAL.
FT DISULFID 260 324 POTENTIAL.
FT DISULFID 369 438 POTENTIAL.
FT DISULFID 478 538 POTENTIAL.
FT CARBOHYD 88 88 N-LINKED (GLCNAAC. . .) (IN ALLOTYPES T61
FT CARBOHYD 108 108 N-LINKED (GLCNAAC. . .) (IN ALLOTYPES T62
AND T63).
FT CARBOHYD 418 418 N-LINKED (GLCNAAC. . .) (IN ALLOTYPES T61).
FT VARIANT 88 88 K -> N (IN ALLOTYPIC T61).
FT VARIANT 94 94 D -> E (IN ALLOTYPIC T61).
FT VARIANT 101 108 TVDQLTQN -> YLNRLSQS (IN ALLOTYPIC T61).
FT VARIANT 110 110 S -> T (IN ALLOTYPIC T63).
SQ SEQUENCE 773 AA; 83886 MW; DF2C44D2F1193C65 CRC64;

Query Match 1.7%; Score 13; DB 1; Length 773;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 727 EPKAKRSKKEA 739
    |||||
Db 736 EPKAKRSKKEA 748

RESULT 6
PIGR_PIG STANDARD; PRT; 102 AA.
AC Q29244;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polymeric-immunoglobulin receptor (Poly-Ig receptor) (PIGR)
DE (Fragment).
GN PIGR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N. A.
RX TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Wintrobe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine CDNA
RT library: analysis of 839 clones.";
RL Mamm. Genome 7:509-517(1996).
RN [2]
RP FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
RN BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
RN TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
RN DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
RN EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
RN TRANSMEMBRANE SEGMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; F14851; CAA23294.1; .

```

DR PROSITE: PS50835; IG-LIKE, PARTIAL.
 KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 102 102
 SO SEQUENCE 102 AA; 11205 MW; 82C915264B1508B8 CRC64;
 Query Match 1.48; Score 11; DB 1; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 234 DLQVLRPEPL 244
 Db 1 DLQVLRPEPL 11
 RESULT 7
 Y82L_METJA STANDARD; PRT; 56 AA.
 AC Q58231;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ0821.
 GN MJ0821.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sulten G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Ulethack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Klenk H.-P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Raebold H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii";
 RL Science 273:1058-1073(1996).
 RN [2]
 RP SIMILARITY.
 RX MEDLINE=97197912; PubMed=9045616;
 RA Koonin E.V.;
 RT "Evidence for a family of archaeal ATPases";
 RL Science 275:1489-1490(1997).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEOAL ATPASE FAMILY. MJ0821, MJ0820
 CC AND MJ0819 RESPECTIVELY REPRESENT THE N-TERMINAL, CENTRAL AND C-
 CC TERMINAL SECTION OF OTHER MEMBERS OF THIS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67526; AAB98833.1; --
 DR PIR: E64402; E64402.
 DR TIGR: MJ0821; --
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 56 AA; 6512 MW; 7E19C78530493060 CRC64;
 Query Match 1.08; Score 8; DB 1; Length 56;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 457 IIEGPNL 464
 I I I I I I I I I I

Db 16 IIEGPNL 23
 RESULT 8
 HIS6_SULSO STANDARD; PRT; 251 AA.
 ID HIS6_SULSO
 AC Q33774;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Imidazole glycerol phosphate synthase subunit hisF (EC 4.1.3.-) (IGP
 DE synthase cyclase subunit) (IGP synthase subunit hisF) (Imgp synthase
 DE subunit hisF) (IGP synthase subunit hisF).
 GN HISF OR S500397 OR C08_052.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=97352708; PubMed=9209067;
 RA Charlebois R.L., Sensen C.W., Doolittle W.F., Brown J.R.;
 RT "Evolutionary analysis of the hisCABDDEHI gene cluster from the
 RT archaeon Sulfolobus solfataricus P2.";
 RL J. Bacteriol. 179:4429-4432(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Welner C.C.-Y., Allard G., Chow C.,
 RA Confalonieri F., Curtis B., Duguet M., Erasuo G., Faguy D.,
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 RA Kushnaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Ragan M.A., Sensen C.W.;
 RT "Gene content and organization of a 281-kbp contig from the genome of
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Awayez M.J., Chan-Welner C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: IGPs catalyzes the conversion of PRFAR and glutamine to
 CC IGP, AICAR and glutamate. The hisF subunit catalyzes the
 CC cyclization activity that produces IGP and AICAR from PRFAR using
 CC the ammonia provided by the hisH subunit (By similarity).
 CC -1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulose-1-yl)imidazole-4-
 CC ylamino)methylideneamino]-1-(5-phosphoribosyl)imidazole-4-
 CC carboxamide + L-glutamine -> imidazole-glycerol phosphate + 5-
 CC aminomidazole-4-carboxamide ribonucleotide + L-glutamate + H(2)O.
 CC -1- PATHWAY: Histidine biosynthesis; fifth step.
 CC -1- SUBUNIT: Heterodimer of hisH and hisF (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE HISA / HISF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U82227; AAB63022.1; -
 DR EMBL: Y18930; CAB57702.1; -
 DR EMBL: AE006689; AAK40908.1; -
 DR PIR: E90206; E90206.
 DR HAMAP: MF_01013; -; 1.
 DR InterPro: IPR003009; FMN_enzyme.
 DR InterPro: IPR006062; His_biosynth.
 DR InterPro: IPR004651; HisF.
 DR Pfam: PF00977; His_biosynth; 1.
 DR TIGRfam: TIGR00735; hisF; 1.
 KW Histidine biosynthesis; Lyase; Complete proteome.
 FT ACT_SITE 11 11 POTENTIAL.
 FT ACT_SITE 130 130 POTENTIAL.
 SO SEQUENCE 251 AA; 26965 MW; 46246934AC1BA850 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 251;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 574 SLAKADAA 581
 DB 215 SLAKADAA 222

RESULT 9:
 T2M1_MICAM STANDARD; PRT; 308 AA.
 ID T2M1_MICAM
 AC P50189;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Type II restriction enzyme MamI (EC 3.1.21.4) (Ecdonuclease MamI)
 DE (R.MamI).
 GN MAMIR.
 OS Microbacterium ammoniaphilum.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Micrococcales; Microbacteriaceae; Microbacterium.
 OX NCBI_TaxID=42460;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15354;
 RX MEDLINE=96257250; PubMed=8654988;
 RA Stiebel H.-M., Seebler S., Jarsch M., Kessler C.;
 RT "Cloning and characterization of the MamI restriction-modification
 RT system from Microbacterium ammoniaphilum in Escherichia coli.";
 RL Gene 172:41-46(1996).
 CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE GATNNNNATC AND
 CC CLEAVES AFTER N-5.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
 CC specific double-stranded fragments with terminal 5'-phosphates.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X79027; CA55648.1; -
 DR PIR: T45133; T45133.
 DR REBASE: 1200; MamI.
 KW Restriction system; Hydrolase; Nuclease; Endonuclease.
 SO SEQUENCE 308 AA; 34293 MW; 53057207BFB1B92 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 308;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 DVVNTLG 287
 DB 200 DVVNTLG 207

RESULT 10
 YD01_METUA STANDARD; PRT; 400 AA.
 ID YD01_METUA
 AC O58697;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ATP-binding protein MJ1301.
 GN MJ1301.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Jal-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 RN [2]
 RP SIMILARITY:
 RC MEDLINE=97197912; PubMed=9045616;
 RA Koornik E.V.;
 RT "Evidence for a family of archaeal ATPases.";
 CC Science 275:1489-1490(1997).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEOAL ATPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U67571; AAB99326.1; -
 DR PIR: D64462; D64462.
 DR TIGR: MJ1301; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002576; Archaeal_ATPase.
 DR Pfam: PF01637; Archaeal_ATPase; 1.
 DR ProDom: PD003808; Archaeal_ATPase; 1.
 DR SMART: SM00382; AAA; 1.
 KW Hypothetical protein; ATP-binding; Complete proteome.
 FT NP_BIND 36 43 ATP (POTENTIAL).
 SO SEQUENCE 400 AA; 46903 MW; EE9460CAD791D049 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 400;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 457 IIEGPNL 464
 DB 23 IIEGPNL 30

RESULT 11
 SYFB_PYRAB STANDARD; PRT; 556 AA.
 ID SYFB_PYRAB
 AC Q9UYX2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
DE tRNA ligase beta chain) (PheRS).
GN PHE1 OR PYRAB13850 OR PAB2427.
OS Pyrococcus abyssi.
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC Pyrococcus
CC NCBI_TaxID=29292;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi."
RL Mol. Microbiol. 47:1495-1512(2003).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
CC CHAIN FAMILY. SUBFAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ248287; CAB50290.1; -
DR PIR: E75049; E75049.
DR HAMAP: MF_00284; 1.
DR InterPro: IPR005147; B5.
DR InterPro: IPR004531; PheT_arch.
DR DR PFam: PF03484; B5; 1.
DR TIGRfams: TIGR00471; PheT_arch; 1.
DR tRNA: tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
FT DOMAIN 135 POLY-ARG.
FT SEQUENCE 556 AA; 63949 MW; 06134413484E60DB CRC64;
SQ
Query Match 1.0%; Score 8; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 396 PLYVDSG 403
Db 200 PLYVDSG 207

```

```

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN
CC FAMILY. SUBFAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE010211; AAL81114.1; -
DR HAMAP: MF_00284; 1.
DR InterPro: IPR005146; B3_4.
DR InterPro: IPR005147; B5.
DR InterPro: IPR004531; PheT_arch.
DR PFam: PF03483; B3_4; 1.
DR PFam: PF03484; B5; 1.
DR TIGRfams: TIGR00471; PheT_arch; 1.
DR tRNA: tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
FT SEQUENCE 556 AA; 63963 MW; CFFB32F6D001432 CRC64;
SQ
Query Match 1.0%; Score 8; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 396 PLYVDSG 403
Db 200 PLYVDSG 207

```

DR EMBL: U04048; AAA50176.1; Interpro: IPR00399; Pyruvate, decarb.
 DR Pfam: PF00205; TPP_enzymes_1.
 DR Pfam: PF02775; TPP_enzymes_C; 1.
 DR Pfam: PF02776; TPP_enzymes_N; 1.
 KW Lyase: Flavoprotein: Thiamine pyrophosphate.
 FT ACT_SITE 50 50
 FT SEQUENCE 563 AA: 58957 MW: 36940C0B33265DBF CRC64:
 Query Match 1.0%; Score 8; DB 1; Length 563;
 Best Local Similarity 100.0%; Pred. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 576 AKADAPP 583
 Db 264 AKADAPP 271
 RESULT 14
 FGFR4_HUMAN STANDARD: PRT; 802 AA.
 ID AC P22455; O43785; Q14309;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Fibroblast growth factor receptor 4 precursor (EC 2.7.1.112) (FGFR-4).
 GN FGFR4 OR JTK2 OR TKF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId:9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE:91224085; PubMed:1709094;
 RA Partanen J.M., Mäkeläe T.P., Eerola E., Korhonen J., Hirvonen H.,
 RA Claesson-Welsh L., Allitalo K.,
 RA "FGFR-4, a novel acidic fibroblast growth factor receptor with a
 RT distinct expression pattern."
 RL EMBO J. 10:1347-1354(1991).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE:Mammary gland;
 RX MEDLINE:93194827; PubMed:7680645;
 RA Ron D., Reich R., Chedid M., Lengel C., Cohen O.E., Chan A.M.,
 RA Neufeld G., Miki T., Tronick S.R.;
 RT "Fibroblast growth factor receptor 4 is a high affinity receptor for
 RT both acidic and basic fibroblast growth factor but not for
 RT keratinocyte growth factor."
 RL J. Biol. Chem. 268:5388-5394(1993).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE:98119018; PubMed:9457674;
 RA Kostzawa M., Müller U.;
 RT "Genomic structure and complete sequence of the human FGFR4 gene."
 RL Mamm. Genome 9:131-135(1998).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Muscle;
 RX MEDLINE:22388257; PubMed:12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Jordan H., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (5)
 RP SEQUENCE OF 609-676 FROM N.A.
 RC TISSUE-Blood;
 RX MEDLINE:91062389; PubMed:2247464;
 RA Partanen J., Mäkeläe T.P., Allitalo R., Lehtoalaaho H., Allitalo K.;
 RT "Putative tyrosine kinases expressed in K-562 human leukemia cells."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8913-8917(1990).
 CC -1- FUNCTION: RECEPTOR FOR ACIDIC FIBROBLAST GROWTH FACTOR. DOES NOT
 CC BIND TO BASIC FIBROBLAST GROWTH FACTOR. BINDS FGF19.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
 CC FAMILY.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC EMBL: X57205; CAA0490.1; -
 CC EMBL: L03840; AAB59389.1; -
 CC EMBL: Y13901; CAA74200.1; -
 CC EMBL: BC011847; AAH11847.1; -
 CC EMBL: M59373; AA63208.1; -
 CC PIR: S15345; TVH0F4.
 CC PDB: 1OCT; 15-JUN-99.
 CC GeneW: HGNC:3691; FGFR4.
 DR MIM: 134935; -
 DR GO: GO:0005887; C: Integral to plasma membrane; TAS.
 DR GO: GO:0005007; F: fibroblast growth factor receptor activity; TAS.
 DR GO: GO:0008343; P: FGF receptor signaling pathway; TAS.
 DR InterPro: IPR007110; 1g-1like.
 DR InterPro: IPR003598; 1g-C2.
 DR InterPro: IPR003006; 1g-MHC.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00047; 1g; 3.
 DR Pfam: PF00069; PKinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00408; IGC2; 3.
 DR SMART: SM00219; TYRKC; 1.
 DR PROSITE: PS50835; 1G-LIKE; 2.
 DR PROSITE: PS50107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 KW Receptor: Glycoprotein: Tyrosine-protein kinase: ATP-binding;
 KW transferase: Phosphorylation; Transmembrane; Immunoglobulin domain;
 KW Repeat: Signal; Polymorphism; 3d-structure.
 FT SIGNAL 25 24
 FT CHAIN 1 802
 FT DOMAIN 25 369
 FT TRANSMEM 370 390
 FT DOMAIN 391 802
 FT DOMAIN 30 128
 FT DOMAIN 152 240
 FT DOMAIN 249 349
 FT DOMAIN 467 755
 FT NP_BIND 473 481
 FT BINDING 503 503
 FT ACT_SITE 612 612
 FT MOD_RES 643 643
 P1: FIBROBLAST GROWTH FACTOR RECEPTOR 4.
 P2: EXTRACELLULAR (POTENTIAL).
 P3: POTENTIAL.
 P4: CYTOPLASMIC (POTENTIAL).
 P5: 1G-LIKE C2-TYPE 1.
 P6: 1G-LIKE C2-TYPE 2.
 P7: 1G-LIKE C2-TYPE 3.
 P8: PROTEIN KINASE.
 P9: ATP (BY SIMILARITY).
 P10: ATP (BY SIMILARITY).
 P11: BY SIMILARITY.
 P12: PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```

FT DISULFID 57 101 POTENTIAL.
FT DISULFID 172 224 POTENTIAL.
FT DISULFID 271 333 POTENTIAL.
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 388 388 G -> R (IN OBSNP:351855).
FT /FTID-VAR.014797.
FT CONFLICT 297 297 D -> V (IN REF. 1).
SQ SEQUENCE 802 AA: 87954 MW: B2B259831BB889F CRC64:

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 802;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
DB 94 EDAGRYLC 101

RESULT 15
FGF4_MOUSE STANDARD: PRT; 808 AA.
ID FGF4_MOUSE
AC 003142;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibroblast growth factor receptor 4 precursor (EC 2.7.1.112) (FGFR-4)
DE (Protein-tyrosine kinase receptor MPK-11).
GN FGF4 OR FGFR-4 OR MPK-11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CD-1; TISSUE=Fetal cerebellum;
RC MEDLINE=92146274; Pubmed=1723680;
RA Stark K.L., McMahon J., McMahon A.P.;
RT "FGFR-4, a new member of the fibroblast growth factor receptor
RT family, expressed in the definitive endoderm and skeletal muscle
RT lineages of the mouse."
RL Development 113:641-651(1991).
RN [2]
RP SEQUENCE OF 620-676 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryonic brain;
RX MEDLINE=93096484; Pubmed=1281307;
RA Gilardi-Hebenstreit P., Nieto M.A., Fraaij M., Mattei M.-G.,
RA Chastier A., Wilkinson D.G., Charney P.;
RT "An Eph-related receptor protein tyrosine kinase gene segmentally
RT expressed in the developing mouse hindbrain."
RL Oncogene 7:2499-2506(1992).
CC -1- FUNCTION: PUTATIVE RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC MAY BE INVOLVED IN THE DEVELOPMENT OF SKELETAL MUSCLE CELL
CC LINEAGES.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING GUT ENDODERM,
CC IN MYOTOMALLY DERIVED SKELETAL MUSCLE, THE ADRENAL CORTEX,
CC KIDNEY AND CONDENSING CARTILAGE. PRESENT IN ADULT LIVER,
CC LUNG AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59927; CAA42551.1; -
DR EMBL: X57236; CAA40512.1; -
DR FIR: S18209; S18209.
DR HSSP: P11362; IFGR.
DR MGD: MGI:95525; Fgf4.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000719; Prot_Kinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00047; Ig_3.
DR Pfam: PF00069; Pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase; 1.
DR SMART: SM00408; Jc2; 3.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS50835; IG_LIKE; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferrase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 808
FT DOMAIN 19 366
FT TRANSMEM 367 387
FT DOMAIN 388 808
FT DOMAIN 53 138
FT DOMAIN 149 237
FT DOMAIN 246 346
FT DOMAIN 464 761
FT NP_BIND 470 478
FT BINDING 509 509
FT ACT_SITE 618 618
FT MOD_RES 649 649
FT DISULFID 53 98
FT DISULFID 169 221
FT DISULFID 268 330
FT CARBOHYD 109 109
FT CARBOHYD 255 255
FT CARBOHYD 287 287
FT CARBOHYD 308 308
FT CARBOHYD 319 319
SQ SEQUENCE 808 AA: 89775 MW: E3F0B5B334E8E6 CRC64:

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 808;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
DB 91 EDAGRYLC 98

```

Search completed: September 9, 2003, 23:17:47
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 9, 2003, 23:14:42 ; Search time 113 Seconds

(without alignments)
1744.710 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764

Sequence: 1 MLFLVLTCLLAVPAISTKTS.....DFLLQSTVAEAADGPOEA 764

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	663	86.8	764	4	081Z77
2	30	3.9	39	4	081Z77
3	22	2.9	360	6	08WJ21
4	17	2.2	758	6	09N2H7
5	10	1.3	82	6	08SQ82
6	10	1.3	217	11	08R544
7	10	1.3	308	17	09HPL3
8	10	1.2	362	12	08JX22
9	9	1.2	362	12	08JX17
10	9	1.2	362	12	08JX17
11	9	1.2	731	6	08SP16
12	9	1.2	733	6	08SQ83
13	8	1.0	48	16	09X7N4
14	8	1.0	122	16	092K99
15	8	1.0	140	6	08S0B8
16	8	1.0	155	16	098F31

17	8	1.0	168	11	08C0K6	08C0K6 mus musculu
18	8	1.0	197	17	057983	057983 pyrococcus
19	8	1.0	283	2	09F212	09F212 streptococc
20	8	1.0	287	11	09D7B8	09D7B8 mus musculu
21	8	1.0	287	12	066659	066659 equine hepr
22	8	1.0	287	16	09KUN5	09KUN5 vibrio chol
23	8	1.0	311	16	08XUL4	08XUL4 ralsionia s
24	8	1.0	496	4	08N311	08N311 homo sapien
25	8	1.0	501	2	093NF2	093NF2 neisseria g
26	8	1.0	509	17	08PRG6	08PRG6 methanosarc
27	8	1.0	562	11	08R113	08R113 mus musculu
28	8	1.0	563	2	09F4L3	09F4L3 pseudomonas
29	8	1.0	565	10	09FHY8	09FHY8 arabidopsis
30	8	1.0	570	11	09DBB9	09DBB9 mus musculu
31	8	1.0	648	5	09V6J4	09V6J4 drosophila
32	8	1.0	654	5	090863	090863 caenorhabdi
33	8	1.0	694	16	P96574	P96574 bacillus su
34	8	1.0	799	11	08C1B8	08C1B8 mus musculu
35	8	1.0	799	11	08C3V5	08C3V5 mus musculu
36	8	1.0	802	4	08TDA0	08TDA0 homo sapien
37	8	1.0	830	5	09BKT7	09BKT7 caenorhabdi
38	8	1.0	1092	11	091ZT0	091ZT0 rattus norv
39	8	1.0	1363	11	091ZT1	091ZT1 rattus norv
40	8	1.0	1553	5	09STR0	09STR0 drosophila
41	8	1.0	2181	5	09VRA6	09VRA6 drosophila
42	7	0.9	36	6	09TRP8	09TRP8 leontopithe
43	7	0.9	58	16	098DT7	098DT7 rhizobium 1
44	7	0.9	63	10	09XIP0	09XIP0 rhizobium 1
45	7	0.9	67	16	08NLT3	08NLT3 corynebacte

ALIGNMENTS

RESULT 1

ID 081Z77 PRELIMINARY: PRT: 764 AA.

AC 081Z77;

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Hepatocellular carcinoma associated protein T86.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Dong X., Pang X., Cheng W.;

RT "Cloning and characterization of hepatocellular carcinoma associated-

RT genes.";

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF272149; AAN65630.1; -

SO SEQUENCE 764 AA; 83283 MW; 927461F8EB305C7 CRC64;

Query Match 86.8%; Score 663; DB 4; Length 764;
Best Local Similarity 99.98; Pred. No. 0;
Matches 763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	1	MLFLVLTCLLAVPAISTKSPICGPEEVNSVEGNSVITCYPTSVNRHTRKWCQGA 60		081Z77	1	MLFLVLTCLLAVPAISTKSPICGPEEVNSVEGNSVITCYPTSVNRHTRKWCQGA 60
DB	1	MLFLVLTCLLAVPAISTKSPICGPEEVNSVEGNSVITCYPTSVNRHTRKWCQGA 60		081Z77	1	MLFLVLTCLLAVPAISTKSPICGPEEVNSVEGNSVITCYPTSVNRHTRKWCQGA 60
OY	61	RGCGITLISSEGVSSKYAGRANLTNPENCTFVNTAQLSDDSGRKCGLSNGLS 120		081Z77	61	RGCGITLISSEGVSSKYAGRANLTNPENCTFVNTAQLSDDSGRKCGLSNGLS 120
DB	61	RGCGITLISSEGVSSKYAGRANLTNPENCTFVNTAQLSDDSGRKCGLSNGLS 120		081Z77	61	RGCGITLISSEGVSSKYAGRANLTNPENCTFVNTAQLSDDSGRKCGLSNGLS 120
OY	121	FDVSLVSVSGGLNDKRYTVTDGRTVTINCPRKTENAKRSLYKQIGLYPVLVIDS 180		081Z77	121	FDVSLVSVSGGLNDKRYTVTDGRTVTINCPRKTENAKRSLYKQIGLYPVLVIDS 180
DB	121	FDVSLVSVSGGLNDKRYTVTDGRTVTINCPRKTENAKRSLYKQIGLYPVLVIDS 180		081Z77	121	FDVSLVSVSGGLNDKRYTVTDGRTVTINCPRKTENAKRSLYKQIGLYPVLVIDS 180
OY	181	GYVPNPTGRRLDIOGTGOLFESVYNOLRLSDAGGYLCAGDGSNNKNADLYLKP 240		081Z77	181	GYVPNPTGRRLDIOGTGOLFESVYNOLRLSDAGGYLCAGDGSNNKNADLYLKP 240
DB	181	GYVPNPTGRRLDIOGTGOLFESVYNOLRLSDAGGYLCAGDGSNNKNADLYLKP 240		081Z77	181	GYVPNPTGRRLDIOGTGOLFESVYNOLRLSDAGGYLCAGDGSNNKNADLYLKP 240

```

Db      181  GYVNPRTYGRIRIDIGTCQQLFSSVYINQLRLSDAGQYLCQAQDSSNKKNAIDLQVLRP 240
Qy      241  EPELVYEDLRGSGVTFFHCALGPVANYAKFLCRSSGENDVYVNLGKRAPAFEGRIILLN 300
Db      241  EPELVYEDLRGSGVTFFHCALGPVANYAKFLCRSSGENDVYVNLGKRAPAFEGRIILLN 300
Qy      301  PODKGSFSSVITTLGRLKEDAGRYLCGASHDGOLOESP1QAMOLFVNEST1PRSPVYK 360
Db      301  PODKGSFSSVITTLGRLKEDAGRYLCGASHDGOLOESP1QAMOLFVNEST1PRSPVYK 360
Qy      361  GYAGSSVAALCPYNRKESKSIKWCMLMEGAONGRCPLVDSEGVAAOYEGRLSLEEPG 420
Db      361  GYAGSSVAALCPYNRKESKSIKWCMLMEGAONGRCPLVDSEGVAAOYEGRLSLEEPG 420
Qy      421  NGFTFYILNQLTSRDAGFYWCMLTNGDTLWRTVEIIEGEPNLKVPGNVTAVLGETLKV 480
Db      421  NGFTFYILNQLTSRDAGFYWCMLTNGDTLWRTVEIIEGEPNLKVPGNVTAVLGETLKV 480
Qy      481  PCHFPCKSSSEYKWKCKMNNQCALPSCDEGPKAFVNCDENSRLVSLTLNLVTRADGC 540
Db      481  PCHFPCKSSSEYKWKCKMNNQCALPSCDEGPKAFVNCDENSRLVSLTLNLVTRADGC 540
Qy      541  WYWCYKQGHFGETAAVVAVERKAAGSRDYSLAKADAPEKVLDSGFREIENKAIQ 600
Db      541  WYWCYKQGHFGETAAVVAVERKAAGSRDYSLAKADAPEKVLDSGFREIENKAIQ 600
Qy      601  DRLFAEERKAVADTRDQADGSRASVDSGSEEGSSRALVSTLVLGLAVAGAVAGV 660
Db      601  DRLFAEERKAVADTRDQADGSRASVDSGSEEGSSRALVSTLVLGLAVAGAVAGV 660
Qy      661  ARARHRKNDVRSIRSIRYTDISMSDFENSRFGANDNMGASSITQETSLGKKEEVAATTE 720
Db      661  ARARHRKNDVRSIRSIRYTDISMSDFENSRFGANDNMGASSITQETSLGKKEEVAATTE 720
Qy      721  STTEKEPKAKRSSKEEAEMAYKDFLLSSTTAAEAQDGPORA 764
Db      721  STTEKEPKAKRSSKEEAEMAYKDFLLSSTTAAEAQDGPORA 764

RESULT 2
O8TE27 PRELIMINARY: PRT: 39 AA.
ID O8TE27;
AC O8TE27;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Polymeric immunoglobulin receptor (Fragment).
GN PIGR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9735866;
RA Pliskurich J.F., Youngman K.R., Phillips K.M., Hempen P.M.,
RA Blanchard M.H., France J.A., Kaetzel C.S.;
RT "Transcriptional regulation of the human polymeric immunoglobulin
RT receptor gene by interferon-gamma.";
RL Mol. Immunol. 34:75-91(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21270451; PubMed=11376148;
RA Bruce S.R., Peterson M.L.;
RT "Multiple features contribute to efficient constitutive splicing of an
RT unusually large exon.";
RL Nucleic Acids Res. 29:2292-2302(2001).
DR EMBL: AF261083; AAF72193.1; -.
KM Receptor.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA: 4272 MW: 3FCBE72AA6467810 CRC64;

```

```

Query Match 3.9%; Score 30; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      319  DAGRYLCGASHDGOLOESPIQAMOLFVNE 348
Db      10  DAGRYLCGASHDGOLOESPIQAMOLFVNE 39

RESULT 3
O8MJ21 PRELIMINARY: PRT: 360 AA.
ID O8MJ21;
AC O8MJ21;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Polymeric immunoglobulin receptor (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Peters I.R., Helps C.R., Hall E.J., Day M.J.;
RT "Cloning and sequencing of canine polymeric immunoglobulin receptor
RT (PIgR) mRNA.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY081057; AAL91653.1; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00409; Ig_1.
KM Receptor.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA: 39685 MW: 32AF72C03FA935CA CRC64;

Query Match 2.9%; Score 22; DB 6; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      104  DSGRYKCGLGINSRGLSPDVS 125
Db      45  DSGRYKCGLGINSRGLSPDVS 66

RESULT 4
O9N2H7 PRELIMINARY: PRT: 758 AA.
ID O9N2H7;
AC O9N2H7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Poly-Ig receptor precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Sone T., Kumura H.;
RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032195; BAA84283.2; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_5.
DR SMART: SM00409; Ig_4.
DR PROSITE: PS50835; IG_LIKE; 2.
KM Receptor; Signal.
FT SIGNAL 1
FT SIGNAL 18
SQ SEQUENCE 758 AA: 83154 MW: D5BBE1A8B082D247 CRC64;

```

Query Match 2.2%; Score 17; DB 6; Length 758;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 116 SRGLSFDVSLVSGPG 132
 DB 116 SRGLSFDVSLVSGPG 132

RESULT 5

OS082 PRELIMINARY; PRT; 82 AA.
 AC OS082;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Polymeric immunoglobulin receptor (Fragment).
 GN PIGR.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;

RN (1)
 RP SEQUENCE FROM N.A.
 RA Pfaffl M.W., Dzidic A., Rojas P.G., Bruckmaier R.M., Schams D.;
 RT "Effects of an induced mammarygenesis and lactogenesis in sheep on the
 RT mRNA expression levels of immune globulin receptors (FcRn, PIGR) and
 RT zona occludens proteins (Occludin, ZO1, ZO2, ZO3)."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ313189; CAC41993.1;
 KM Receptor.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 82 AA; 8846 MW; D6040018954A50AA CRC64;

Query Match 1.3%; Score 10; DB 6; Length 82;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 145 GRTVINCPE 154
 DB 23 GRTVINCPE 32

RESULT 6

OS0844 PRELIMINARY; PRT; 217 AA.
 AC OS0844;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Polymeric immunoglobulin receptor (Fragment).
 GN PIGR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN-129/Sv;
 RC MEDLINE-99377172;
 RX MEDLINE-99377172;
 RA Bruce S.R., Kaetzel C.S., Peterson M.L.;
 RT "Cryptic intron activation within the large exon of the mouse
 RT polymeric immunoglobulin receptor gene: cryptic splice sites
 RT correspond to protein domain boundaries.";
 RL Nucleic Acids Res. 27:3446-3454(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA STRAIN-129/Sv;
 RC MEDLINE-21270451; PubMed-11376148;
 RX Bruce S.R., Peterson M.L.;

RT "Multiple features contribute to efficient constitutive splicing of an
 RT unusually large exon.";
 RL Nucleic Acids Res. 29:2292-2302(2001).
 DR EMBL: AF261084; AAF72194.1;
 DR MGD: MG1:103080; PIGR.
 DR InterPro: IPR003599; IG.
 DR SMART: SM00409; IG; 2.
 DR Receptor.
 KM Receptor.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 217 AA; 24352 MW; B579C870ACFF180FB CRC64;

Query Match 1.3%; Score 10; DB 11; Length 217;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 312 ITGLKEDAG 321
 DB 181 ITGLKEDAG 190

RESULT 7

OS08PL3 PRELIMINARY; PRT; 308 AA.
 AC OS08PL3;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Vng1578h.
 GN VNG1578H.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20504483; PubMed-11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasry S.R., Baliga N.S., Thorsson V., Shroga J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weltl R., Goo Y.A.,
 RA Lettner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL: AE005070; AAG19854.1;
 DR InterPro: IPR004485; CblB.
 DR Pfam: PF03186; CblB_CblB; 1.
 KM Complete proteome.
 SQ SEQUENCE 308 AA; 30446 MW; 3FA8778B472D7ECB CRC64;

Query Match 1.3%; Score 10; DB 17; Length 308;
 Best Local Similarity 100.0%; Pred. No. 0.34;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 652 AVGAVAGVA 661
 DB 5 AVGAVAGVA 14

RESULT 8

OS08X22 PRELIMINARY; PRT; 362 AA.
 AC OS08X22;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 GN P1.
 OS Turnip mosaic virus (strain Japanese) (TuMV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;

OC Polyvirus.
 OX NCBI_TaxID-12230;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AL;
 RX MEDLINE-22024854; PubMed-12029167;
 RA Ohshima K., Yamaguchi Y., Hirota R., Hamamoto T., Tomimura K., Tan Z.,
 RA Sano T., Azuhata F., Walsh J.A., Fletcher J., Chen J., Gera A.,
 RA Gibbs A.;
 RT "Molecular evolution of Turnip mosaic virus: evidence of host
 adaptation, genetic recombination and geographical spread.";
 RL J. Gen. Virol. 83:1511-1521(2002).
 DR EMBL: AB076441; BAC02777.1;
 DR InterPro: IPR002540; Poly_P1.
 DR Pfam: PF01577; Poly_P1; 1.
 FT NON_TER 362
 SQ SEQUENCE 362 AA; 40817 MW; FD2B9F108A1870E8 CRC64;

Query Match 1.2%; Score 9; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 VVKGVAGSS 366
 Db 63 VVKGVAGSS 71

RESULT 9
 O8JX17 PRELIMINARY; PRT; 362 AA.
 AC O8JX17;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 GN P1.
 OS Turnip mosaic virus (strain Japanese) (TUNV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
 OC Polyvirus.
 OX NCBI_TaxID-12230;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1S1;
 RX MEDLINE-22024854; PubMed-12029167;
 RA Ohshima K., Yamaguchi Y., Hirota R., Hamamoto T., Tomimura K., Tan Z.,
 RA Sano T., Azuhata F., Walsh J.A., Fletcher J., Chen J., Gera A.,
 RA Gibbs A.;
 RT "Molecular evolution of Turnip mosaic virus: evidence of host
 adaptation, genetic recombination and geographical spread.";
 RL J. Gen. Virol. 83:1511-1521(2002).
 DR EMBL: AB076446; BAC02782.1;
 DR InterPro: IPR002540; Poly_P1.
 DR Pfam: PF01577; Poly_P1; 1.
 FT NON_TER 362
 SQ SEQUENCE 362 AA; 40466 MW; ACE990B0B9CE8A44 CRC64;

Query Match 1.2%; Score 9; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 VVKGVAGSS 366
 Db 63 VVKGVAGSS 71

RESULT 10
 O8JWY1 PRELIMINARY; PRT; 362 AA.
 AC O8JWY1;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Polyprotein (Fragment).

GN P1.
 OS Turnip mosaic virus (strain Japanese) (TUNV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
 OC Polyvirus.
 OX NCBI_TaxID-12230;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GBR7;
 RX MEDLINE-22024854; PubMed-12029167;
 RA Ohshima K., Yamaguchi Y., Hirota R., Hamamoto T., Tomimura K., Tan Z.,
 RA Sano T., Azuhata F., Walsh J.A., Fletcher J., Chen J., Gera A.,
 RA Gibbs A.;
 RT "Molecular evolution of Turnip mosaic virus: evidence of host
 adaptation, genetic recombination and geographical spread.";
 RL J. Gen. Virol. 83:1511-1521(2002).
 DR EMBL: AB076482; BAC02818.1;
 DR InterPro: IPR002540; Poly_P1.
 DR Pfam: PF01577; Poly_P1; 1.
 FT NON_TER 362
 SQ SEQUENCE 362 AA; 40267 MW; B991D09CA4C85F25 CRC64;

Query Match 1.2%; Score 9; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 VVKGVAGSS 366
 Db 63 VVKGVAGSS 71

RESULT 11
 O8SPI6 PRELIMINARY; PRT; 731 AA.
 AC O8SPI6;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Polymeric immunoglobulin receptor.
 GN PIGR.
 OS Macropus eugenii (Tamar wallaby).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID-9315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Mammary lymph node;
 RA Taylor C.L., Harrison G.A.;
 RT "cDNA cloning of the polymeric immunoglobulin receptor of the
 marsupial Macropus eugenii (Tamar wallaby).";
 RT Submitted (OCT-2000) to the EMBL/Genbank/DBD databases.
 RL J. Gen. Virol. 83:17205; AAK69593.1;
 DR EMBL: AF317205; AAK69593.1;
 DR InterPro: IPR003599; Ig_
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_5.
 DR SMART: SM00409; Ig_5.
 DR PROSITE: P550835; Ig-LIKE; 3.
 DR SMART: SM00409; Ig_5.
 DR PROSITE: P550835; Ig-LIKE; 3.
 KW Receptor.
 SQ SEQUENCE 731 AA; 80253 MW; 91BD3BC3478BC17 CRC64;

Query Match 1.2%; Score 9; DB 6; Length 731;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 FTVINQLT 432
 Db 420 FTVINQLT 428

RESULT 12
 O8SOB3 PRELIMINARY; PRT; 733 AA.
 ID O8SOB3
 AC O8SOB3;

DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Polymeric-Immunoglobulin receptor.
 GN PI3R.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
 NCBI_TaxID=9337;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RA Adamski F.M., Demmer J.;
 RT "Cloning and characterization of PI3R and J chain of the marsupial,
 RT Trichosurus vulpecula (brush-tailed possum).";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF091137; AAD1688.1;
 DR InterPro: IPR003599; IG_1.
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam: PF00047; IG_5.
 DR SMART: SM00409; IG_5.
 DR PROSITE: PS50835; IG_LIKE; 3.
 SK RECEPTOR
 SQ SEQUENCE 733 AA; 80300 MW; 8168872D8AC14A5A CRC64;
 Query Match 1.0%; Score 8; DB 6; Length 733;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 424 FTVILNQLT 432
 DB 422 FTVILNQLT 430
 RESULT 13
 Q9X7N4 PRELIMINARY; PRT; 48 AA.
 AC Q9X7N4;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Small hypothetical protein.
 GN SC06728 OR SC5F2A.11.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=2196410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: AL939129; CAB40678.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 48 AA; 4983 MW; 93710C77F7944475 CRC64;
 Query Match 1.0%; Score 8; DB 16; Length 48;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 648 GLVLA VCA 655

DB 24 GLVLA VCA 31
 RESULT 14
 Q92K99 PRELIMINARY; PRT; 122 AA.
 AC Q92K99;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Hypothetical protein R01815.
 GN R01815 OR SMC00191.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Rampsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591788; CAC46394.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 122 AA; 12774 MW; F3C77A35579D8C21 CRC64;
 Query Match 1.0%; Score 8; DB 16; Length 122;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 651 LAVGAVAV 658
 DB 79 LAVGAVAV 86
 RESULT 15
 Q8S0D8 PRELIMINARY; PRT; 140 AA.
 AC Q8S0D8;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Vascular endothelial growth factor receptor 3 (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoon Y.-S., Murayama T., Isner J.M.;
 RT "Favorable effect of VEGF-C Gene Transfer on Lymphedema.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF453570; AAL78955.1;
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003598; IG_C2.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam: PF00047; IG_1.
 DR SMART: SM00408; IGC2; 1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 KW Immunoglobulin domain; Receptor.
 FT NON_TER 1
 FT TER 1
 SQ SEQUENCE 140 AA; 15340 MW; C95ADA0160981E11 CRC64;
 Query Match 1.0%; Score 8; DB 6; Length 140;
 Best Local Similarity 100.0%; Pred. No. 20;

Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	318	EDAGRYLC	325						
Db	44	EDAGRYLC	51						

Search completed: September 9, 2003, 23:20:39
Job time : 116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:06:57 ; Search time 84 Seconds

(without alignments)
1443.656 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764
Sequence: 1 MLFLVNLCLLAVFPAISTKS.....DFLLQSSTVAAEAQDGPQEA 764

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	22	AA65695
2	764	100.0	764	23	AA65711
3	764	100.0	764	23	ABJ04350
4	746	97.6	746	17	AAW03178
5	607	79.5	607	20	AAW95601
6	605	79.2	607	20	AAV34099
7	605	79.2	607	23	AAW47867
8	254	33.2	272	20	AAV73981
9	243	31.8	243	24	ABP55307

10	243	31.8	243	24	ABP55315	Human polyimmunog1
11	94	12.3	94	24	ABP55311	Human polyimmunog1
12	90	11.8	90	22	AA65712	Human polymeric im
13	61	8.0	61	19	AAW43098	Polymeric immunog1
14	61	8.0	61	19	AAW43099	Polymeric immunog1
15	60	7.9	60	20	AAW99073	Immunoglobulin 11k
16	55	7.2	243	23	AAE29187	Cynomolgus monkey
17	55	7.2	243	24	ABP55308	Simian polyimmunog
18	55	7.2	243	24	ABP55316	Simian polyimmunog
19	55	7.2	243	24	ABP55317	Simian polyimmunog
20	50	6.5	70	21	AAW53678	Human colon cancer
21	44	5.8	243	24	ABP55306	Polyimmunoglobulin
22	43	5.6	43	20	AAW9078	Immunoglobulin 11k
23	31	4.1	31	19	AAW43095	Human polymeric im
24	31	4.1	31	22	AA65713	Human polymeric im
25	31	4.1	243	24	ABP55314	Polyimmunoglobulin
26	29	3.8	46	20	AAW99077	Immunoglobulin 11k
27	27	3.5	57	20	AAW99072	Immunoglobulin 11k
28	22	2.9	94	24	ABP55310	Polyimmunoglobulin
29	21	2.7	757	17	AAW03179	Bovine poly-immuno
30	21	2.7	757	22	AA65696	Bovine poly-immuno
31	20	2.6	94	24	ABP55312	Simian polyimmunog
32	19	2.5	19	20	AAW93888	Bifidobacterium bi
33	19	2.5	40	19	AAW43097	Rabbit polymeric 1
34	18	2.4	18	23	ABG60662	Polyimmunoglobulin
35	18	2.4	23	23	ABG60663	Polyimmunoglobulin
36	18	2.4	24	23	ABG60664	Polyimmunoglobulin
37	17	2.2	769	17	AAW03181	Rat poly-immunoglo
38	17	2.2	769	22	AA65697	Rat polymeric immu
39	17	2.2	771	23	ABG60638	Chimeric polyimmun
40	16	2.1	16	19	AAW64615	Human polyimmunog1
41	16	2.1	16	19	AAW61592	Polyimmunoglobulin
42	16	2.1	16	20	AAW85768	Polyimmunoglobulin
43	16	2.1	16	23	ABG94831	Human polyimmunog1
44	16	2.1	16	23	ABG68266	Targeting molecul
45	16	2.1	771	17	AAW03180	Mouse poly-immunog

ALIGNMENTS

RESULT 1	
AA65695	AA65695 standard; protein; 764 AA.
XX	XX
AC	AA65695;
XX	XX
DT	07-JAN-2002 (first entry)
XX	XX
DE	Human polymeric immunoglobulin receptor (p1gR) sequence.
XX	XX
KW	Polymeric immunoglobulin receptor; p1gR; ligand; therapeutic;
KW	Carcinoma diagnosis; veterinary; human.
XX	XX
OS	Homo sapiens.
XX	XX
FT	FT
Key	Location/Qualifiers
FT	Peptide
FT	487..603
FT	/note- "peptide to which a ligand binds to (claim 8)"
FT	487..607
FT	/note- "peptide to which a ligand binds to"
FT	487..611
FT	/note- "peptide to which a ligand binds to"
FT	487..615
FT	/note- "peptide to which a ligand binds to"
FT	487..618
FT	/note- "peptide to which a ligand binds to"
FT	520..607
FT	/note- "peptide to which a ligand binds to"
FT	520..611
FT	/note- "peptide to which a ligand binds to"
FT	520..615
FT	/note- "peptide to which a ligand binds to"

KW Polymetric immunoglobulin receptor; pigr: ligand; therapeutic;
 KW carcinoma diagnosis; veterinary; human.
 XX Homo sapiens.
 XX WO200172846-A2.
 PD 04-OCT-2001.
 XX 26-MAR-2001; 2001WO-US09699.
 XX 27-MAR-2000; 2000US-192197P.
 PR 27-MAR-2000; 2000US-192198P.
 XX (REGC) UNIV CALIFORNIA.
 PA Mostov KE, Chapin SJ, Richman-Eisenstat J;
 PI WPI; 2001-611619/70.
 XX
 XX New ligands binding to a specific region of a polymetric immunoglobulin
 PT receptor, useful for transporting therapeutic or diagnostic
 PT compositions into or across cells expressing pigr e.g. in drug delivery
 XX
 XX Disclosure; Fig 2; 102pp; English.
 PS
 XX The invention provides ligands that bind specifically to a region of an
 CC animal cell polymetric immunoglobulin receptor (pigr). The pigr cleaves
 CC to produce a stalk region remaining attached to the cell and a secretory
 CC component existing in the organ of interest in several forms. The ligands
 CC do not bind to the stalk or the most abundant form of the secretory
 CC component present in the organ under physiological conditions. The
 CC ligands are useful for transporting therapeutic or diagnostic
 CC compositions into or across cells expressing pigr, useful to introduce
 CC or transport ligands such as antibodies and/or to deliver biologically
 CC active components such as proteins, nucleic acids or detectable labels.
 CC They are used to deliver therapeutic compositions to mucosal surfaces
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.
 CC They are also useful to label cells expressing pigr, e.g. to distinguish
 CC epithelial cells from a mixed cell population in pathology studies or to
 CC aid in carcinoma diagnosis (since pigr expression is reduced in
 CC carcinomas relative to normal epithelium). They can also be used to
 CC deliver veterinary compositions, especially in mammals such as farm,
 CC domestic or wild mammals or birds e.g. birds reared for human
 CC consumption. The present sequence represents a human pigr sequence.
 CC
 XX
 SO Sequence 764 AA:
 Query Match 100.0%; Score 764; DB 22; Length 764;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 301 PODKGSFVSIVITGLRKEDAGRYLCAHSDGQLQESPIQAMOLFVNEESTIPRSPTVVK 360
 DB 301 PODKGSFVSIVITGLRKEDAGRYLCAHSDGQLQESPIQAMOLFVNEESTIPRSPTVVK 360
 OY 361 GVAGSSVAVLCPRNRESKSIKYWCIMEGANORCLVDSEGVAAQVREGRLSLEBPQ 420
 DB 361 GVAGSSVAVLCPRNRESKSIKYWCIMEGANORCLVDSEGVAAQVREGRLSLEBPQ 420
 OY 421 NGFTVILNQLTSRDAGFYWCLTNGDTLMTFTVEIKIIEGEPILKYPGNVAVLGETLKY 480
 DB 421 NGFTVILNQLTSRDAGFYWCLTNGDTLMTFTVEIKIIEGEPILKYPGNVAVLGETLKY 480
 OY 481 PCHEPCKFSSEYKWKMNNTGQALPSODEGSPKAFVNCDENSRLVSLTLNVTBDEG 540
 DB 481 PCHEPCKFSSEYKWKMNNTGQALPSODEGSPKAFVNCDENSRLVSLTLNVTBDEG 540
 OY 541 WYMGYKQGHFVGETAAVYAAVEERKAAGSRDVLAKADAPDEKVLDSGFREIENKATQ 600
 DB 541 WYMGYKQGHFVGETAAVYAAVEERKAAGSRDVLAKADAPDEKVLDSGFREIENKATQ 600
 OY 601 DPLFAEEKAVADTRQADGSRASVDSGSEEGCGSSRALVSTLPVLGLAVGAAGV 660
 DB 601 DPLFAEEKAVADTRQADGSRASVDSGSEEGCGSSRALVSTLPVLGLAVGAAGV 660
 OY 661 ARARHKNDRVRSIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKEEYATTE 720
 DB 661 ARARHKNDRVRSIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKEEYATTE 720
 OY 721 STTETKEPKKAKRSKEEAEMAKDVLLOSSTYAAADQCPQEA 764
 DB 721 STTETKEPKKAKRSKEEAEMAKDVLLOSSTYAAADQCPQEA 764
 RESULT 3
 ID ABJ04350 standard; Protein: 764 AA.
 XX ABJ04350;
 AC
 XX 11-OCT-2002 (first entry)
 DT
 XX Human colon specific protein SEQ ID NO: 124.
 DE
 XX Human: colon specific gene; colon specific protein; colon cancer;
 KW colorectal cancer; colon disease; cytostatic; gene therapy.
 CC
 OS Homo sapiens.
 XX
 XX WO200242460-A2.
 PN 30-MAY-2002.
 XX 21-NOV-2001; 2001WO-US43611.
 PF 22-NOV-2000; 2000US-252505P.
 PR (DIAD-) DIADEXUS INC.
 PA Machina RA, Reclipon H, Plura J, Ghosh MG, Sun Y, Liu C;
 PI WPI; 2002-583378/62.
 DR
 XX Novel colon specific polypeptides and polynucleotides useful for
 PT detecting, diagnosing, monitoring, treating, staging and predicting
 PT cancers in humans having cancer and non-cancerous colon disease
 PS Claim 11; Page 225-228; 228pp; English.
 XX The present invention provides protein and coding sequences of human
 CC colon specific genes and proteins. These can be used in the treatment of
 CC colonic diseases, including colon and colorectal cancers. The present
 CC sequence is a protein of the invention.

Db 1 KSPIFGPEEVNSVSGSVSTICYPTTYSVNRHTRKWCROGARGCITLISSEGYSSKY 60
 QY 79 AGRANLTNPENGTFVFNIAQLSODDSGRYKCGLGINSRGLSPFVLSVOGPGLLNDR 138
 Db 61 AGRANLTNPENGTFVFNIAQLSODDSGRYKCGLGINSRGLSPFVLSVOGPGLLNDR 120
 QY 139 VYTVDLGRTVTINCPKTEMAOKRKSLYKQIGLYPVLVIDSGVNPNTGRTLRDIOGT 198
 Db 121 VYTVDLGRTVTINCPKTEMAOKRKSLYKQIGLYPVLVIDSGVNPNTGRTLRDIOGT 180
 QY 199 GOLLFSVYNOLRLSDAGQYLCQAGDSDNSNKNADLOYLKPEPELYEDLKGSVTFHCA 258
 Db 181 GOLLFSVYNOLRLSDAGQYLCQAGDSDNSNKNADLOYLKPEPELYEDLKGSVTFHCA 240
 QY 259 LGPEVANNAKFLCROSSGENDVYVNTLGKRAPAFEGRTILNPODDGSEFSVYITLGRKE 318
 Db 241 LGPEVANNAKFLCROSSGENDVYVNTLGKRAPAFEGRTILNPODDGSEFSVYITLGRKE 300
 QY 319 DAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVKGAGSSVAVALCPYNRKES 378
 Db 301 DAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVKGAGSSVAVALCPYNRKES 360
 QY 379 KSITKWCLMEGAQNGRCPLVDSEGWKAQYEGRLSLLEPENGFTVILNOLTSRDAGF 438
 Db 361 KSITKWCLMEGAQNGRCPLVDSEGWKAQYEGRLSLLEPENGFTVILNOLTSRDAGF 420
 QY 439 VYCLTNGDTLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKYPCHPKCFSSYEKYMCKM 498
 Db 421 VYCLTNGDTLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKYPCHPKCFSSYEKYMCKM 480
 QY 499 NNTGCOALPDSODEGSKAFVNCDENSRLVSLTLNLVTRADEGMYWGVKQGHFYGETAAV 558
 Db 481 NNTGCOALPDSODEGSKAFVNCDENSRLVSLTLNLVTRADEGMYWGVKQGHFYGETAAV 540
 QY 559 YVAVERKAAGSRDVSIAKADAAPDEKVIDSGPRETENKAIODPRFAEKKAAVDTRDQA 618
 Db 541 YVAVERKAAGSRDVSIAKADAAPDEKVIDSGPRETENKAIODPRFAEKKAAVDTRDQA 600
 QY 619 DGSRAVSVDGSEEDGSSRALVSTLPLGLVLAAGAVAVGARAHRKRVDRVSTRSYR 678
 Db 601 DGSRAVSVDGSEEDGSSRALVSTLPLGLVLAAGAVAVGARAHRKRVDRVSTRSYR 660
 QY 679 TDISMSDFENSRFCANDNMGASSTJOETSLGCKEEFVATTESTETKEPKKAKSSKEE 738
 Db 661 TDISMSDFENSRFCANDNMGASSTJOETSLGCKEEFVATTESTETKEPKKAKSSKEE 720
 QY 739 AEMAYKDFLOSSTVAAPADGPOEA 764
 Db 721 AEMAYKDFLOSSTVAAPADGPOEA 746
 RESULT 5
 AAM95601
 ID AAM95601 standard: Protein: 607 AA.
 AC AAM95601:
 DT 08-JUN-1999 (first entry)
 DE Human secretory Immunoglobulin A component.
 KW Immunoglobulin A: secretory; component: IGA; human; treatment:
 KW prevention; infection: HIV; AIDS; cold; flu; virus;
 KW human immunodeficiency virus; respiratory syncytial virus.
 OS Homo sapiens.
 XX
 XX W09857993-AL.
 XX
 XX 23-DEC-1998.
 XX PD
 XX PF 10-JUN-1998: 98WO-US11975.
 XX

PR 19-JUN-1997: 97US-0050969.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Chintalacharuvu KR, Morrison SL;
 PI
 XX
 XX WPI: 1999-080950/07.
 DR N-PSDB: AAX07407.
 DR
 XX
 PT Producing secretory immunoglobulin in single cells - useful to
 PT produce commercial quantities of secretory immunoglobulin to prevent
 PT or treat infections
 PS
 XX Disclosure: Pages 22-24; 39pp; English.
 CC
 CC The sequence is that of the secretory component of human secretory
 CC immunoglobulin A (siga). It can be used as part of a method for
 CC the production of sig molecules. This method is useful for
 CC producing commercial quantities of sig (especially siga) to treat
 CC or prevent infections. In particular, siga produced by the method
 CC can be used to prevent or treat infections in mammals, birds or
 CC fish; especially systemic infections or infections at a mucosal
 CC surface. It is especially useful to prevent or treat infection
 CC with human immunodeficiency virus (HIV), respiratory syncytial
 CC virus, flu virus or cold virus. The method allows production of
 CC commercial quantities of sig molecules for therapeutic use, not
 CC previously possible: production using non-plant cells and a
 CC single cell type is more efficient than a previous multi-step
 CC process of fusing recombinant plant cells, and avoids alterations
 CC of the sig by plant cells. Siga molecules are more stable
 CC and resistant to proteolysis than previously used Iga molecules,
 CC and can be administered to prevent as well as to treat infections,
 CC unlike e.g. Igg and Igm molecules.
 CC
 XX
 SO Sequence 607 AA.
 Query Match 79.5%; Score 607; DB 20; Length 607;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLFLVTLCLAAFPAPISRTKSPIFGPEEVNSVSGSVSTICYPTTYSVNRHTRKWCROGA 60
 Db 1 MLFLVTLCLAAFPAPISRTKSPIFGPEEVNSVSGSVSTICYPTTYSVNRHTRKWCROGA 60
 QY 61 RGGCTLLISSEGYSSKXAGRANLTNPENGTFVFNIAQLSODDSGRYKCGLGINSRGLS 120
 Db 61 RGGCTLLISSEGYSSKXAGRANLTNPENGTFVFNIAQLSODDSGRYKCGLGINSRGLS 120
 QY 121 FDVLSLEVSOGPGLNDTAVYTVDLGRTVTINCPKTEMAOKRKSLYKQIGLYPVLVIDSS 180
 Db 121 FDVLSLEVSOGPGLNDTAVYTVDLGRTVTINCPKTEMAOKRKSLYKQIGLYPVLVIDSS 180
 QY 181 GYVNPNTGRTLRDIOGTGOLLFSVYNOLRLSDAGQYLCQAGDSDNSNKNADLOYLKLP 240
 Db 181 GYVNPNTGRTLRDIOGTGOLLFSVYNOLRLSDAGQYLCQAGDSDNSNKNADLOYLKLP 240
 QY 241 EPELYEDLKGSVTFHCAALGPEVANNAKFLCROSSGENDVYVNTLGKRAPAFEGRTILN 300
 Db 241 EPELYEDLKGSVTFHCAALGPEVANNAKFLCROSSGENDVYVNTLGKRAPAFEGRTILN 300
 QY 301 PODKDGSPSVYITGRLKEDAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVVK 360
 Db 301 PODKDGSPSVYITGRLKEDAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVVK 360
 QY 361 GYAGSSVAVILCPYNRKESKSIKWKLMBSAQNGRCPLVDSEGWKAQYEGRLSLLEPENG 420
 Db 361 GYAGSSVAVILCPYNRKESKSIKWKLMBSAQNGRCPLVDSEGWKAQYEGRLSLLEPENG 420
 QY 421 NGTFVILNQLTSRDAGFYWCLJTNGDTLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKY 480
 Db 421 NGTFVILNQLTSRDAGFYWCLJTNGDTLMRTTVEIKIIEGEPNLKVGNNTAVAGETLKY 480
 QY 481 PCHPCKCFSSYEKYMCKMNTGCOALPDSODEGSKAFVNCDENSRLVSLTLNLVTRADEG 540

```

Db      481  |||||||
          PCHPCKFSSEYKWKCKNNNTGCOALPSODEGPKAAVNCDENSRLVSLTLNLTTRADG 540
Qy      541  WYMGVKGQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIQ 600
          |||||||
Db      541  WYMGVKGQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIQ 600
          |||||||
Qy      601  DPRLFAE 607
          |||||||
Db      601  DPRLFAE 607

```

RESULT 6

AAV34099
ID AAV34099 standard; Protein: 607 AA.

AC AAV34099;

DT 20-DEC-1999 (first entry)

DE Partial amino acid sequence of plasmid pSHUSC.

KW Multimeric protein; immunoglobulin; receptor-ligand complex;

KV hetero-dimeric receptor; trimeric G protein; transgenic.

OS Synthetic.

PN MO9949024-A2.

PD 30-SEP-1999.

PF 24-MAR-1999; 99MO-US06506.

PR 25-MAR-1998; 98US-0079249.

PA (PLAN-) PLANET BIOTECHNOLOGY INC.

PI Wycoff KL, Jaiswal SK;

DR WPI: 1999-580446/49.

DR N-PSDB: AA222290.

PT Producing heterologous multimeric proteins in plants, transformed with several plasmids expressing polypeptide components, particularly for immunoglobulins

PS Example 1; Fig 8; 42pp: English.

XX The invention relates to a method for producing heterologous, multimeric proteins in plant cells. The method comprises: (a) transforming the cells with several naked plasmids each encoding some, but not all, of the polypeptide components of the multimeric proteins, and together providing all the polypeptide components; and (b) culturing the cells. The method is used to produce biologically active multimeric proteins, particularly immunoglobulins, receptor-ligand complexes, homo- or hetero-dimeric receptors, or trimeric G proteins. This method provides properly associated and assembled multimeric proteins in a fast and efficient process, without the need to cross plants expressing single component of the protein. Transgenic plants containing adjacent and stably integrated CC plasmids, and their progeny can also express the multimeric proteins. The CC present sequence represents the partial amino acid sequence of the plasmid pSHUSC.

CC Sequence 607 AA;

SQ XX

Query Match 79.2%; Score 605; DB 20; Length 607;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      3  LFLVLTCLAVPAISTKSPFGPEEVNSVSGNSITCYPPTSVNRHTRKRYMCRGARG 62
          |||||||
Db      3  LFLVLTCLAVPAISTKSPFGPEEVNSVSGNSITCYPPTSVNRHTRKRYMCRGARG 62

```

```

Qy      63  GCITLLISEGYVSSKTAGRANLTFPENGTFVYVNTIQLSDDSGRYKCGLGINSRLSFD 122
          |||||||
Db      63  GCITLLISEGYVSSKTAGRANLTFPENGTFVYVNTIQLSDDSGRYKCGLGINSRLSFD 122
          |||||||
Qy      123  VSLLEVSGPCLLNDTFVYVTDLGRVTYINCFKTEENOKRSLSYKQIGLVPVIVISSG 182
          |||||||
Db      123  VSLLEVSGPCLLNDTFVYVTDLGRVTYINCFKTEENOKRSLSYKQIGLVPVIVISSG 182
          |||||||
Qy      183  VNPNTYGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCAQADDNSNKKKADLQVLKPEP 242
          |||||||
Db      183  VNPNTYGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCAQADDNSNKKKADLQVLKPEP 242
          |||||||
Qy      243  ELVYEDLRGSVTFHICALGPEVAVAAFLCQSSGECNDVYVNTLGRRAAFEGRIILLNQ 302
          |||||||
Db      243  ELVYEDLRGSVTFHICALGPEVAVAAFLCQSSGECNDVYVNTLGRRAAFEGRIILLNQ 302
          |||||||
Qy      303  DKDGFSFVYTTGRLKEDAGRYLCGASHDGLQSGSPLOAMQLFVNEESTIPRSPVYKGV 362
          |||||||
Db      303  DKDGFSFVYTTGRLKEDAGRYLCGASHDGLQSGSPLOAMQLFVNEESTIPRSPVYKGV 362
          |||||||
Qy      363  AGSSVAVLCPYNNKESKSIKYMCLMGAQNGRCPLLVDSGMYKAQYEGRLSLLEPGNG 422
          |||||||
Db      363  AGSSVAVLCPYNNKESKSIKYMCLMGAQNGRCPLLVDSGMYKAQYEGRLSLLEPGNG 422
          |||||||
Qy      423  TFTVILNQLTSRDAGTYWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKYPC 482
          |||||||
Db      423  TFTVILNQLTSRDAGTYWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKYPC 482
          |||||||
Qy      483  HPPCKFSSEYKWKCKNNNTGCOALPSODEGPKAAVNCDENSRLVSLTLNLTTRADG 542
          |||||||
Db      483  HPPCKFSSEYKWKCKNNNTGCOALPSODEGPKAAVNCDENSRLVSLTLNLTTRADG 542
          |||||||
Qy      543  WCGVKGQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIQDP 602
          |||||||
Db      543  WCGVKGQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIQDP 602
          |||||||
Qy      603  RLFAE 607
          |||||||
Db      603  RLFAE 607

```

RESULT 7

AA47867
ID AA47867 standard; Protein: 607 AA.

AC AA47867;

DT 22-FEB-2002 (first entry)

DE Human secretory component.

KW Human; immunoadhesin; intercellular adhesion molecule; ICAM-1;

KV human rhinovirus; immunoglobulin heavy chain; J chain; HRV; common cold;

KW transgenic plant.

OS Homo sapiens.

PN WO200183529-A2.

PD 08-NOV-2001.

PF 28-APR-2001; 2001MO-US13932.

PR 28-APR-2000; 2000US-200298P.

PA (PLAN-) PLANET BIOTECHNOLOGY INC.

PI Larrick JW, Wycoff KL;

DR WPI: 2002-041481/05.
DR N-PSDB: ABA05260.
PT Immunoadhesin for treating human rhinovirus infection comprises

chimeric intercellular adhesion molecule-1, and optionally a J chain and secretory component in association -

The invention relates to an immunoadhesin comprising:
(a) a chimeric intercellular adhesion molecule (ICAM)-1 comprising a rhinovirus receptor protein linked to at least a portion of an immunoglobulin heavy chain; and

(b) optionally a J chain and secretory component associated with the chimeric ICAM-1 molecule. The immunoadhesin has plant-specific glycosylation and virucide activity. The immunoadhesin is useful for reducing infection by human rhinovirus (HRV) and hence the initiation or spread of the common cold by HRV. The immunoadhesin binds to HRV and reduces its infectivity, competing with cell surface ICAM-1 for binding sites, interfering with virus entry or uncoating and directing premature release of viral RNA and formation of empty capsids. Expression of the immunoadhesin in plants would be tetrameric, rather than dimeric. Immunoadhesins having multiple binding sites have a higher effective affinity for the virus, thereby increasing the effectiveness of the immunoadhesin. Association of secretory component and immunoglobulin J chain increases the stability of the immunoadhesin in the mucosal environment. Production is significantly less expensive in plants than in animal cell culture and production in plants is safer for human use, since plants are not known to harbor any animal viruses. The present sequence is that of the human secretory component expressed from the plasmid pSHuSc₂ of the invention.

SQ Sequence 607 AA;

Query Match	79.2%	Score 605;	DB 23;	Length 607;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 605; Conservative	0;	Mismatches	0;	Gaps 0

Qy	J	LFPLVTCLLAVFPAISIKSPITGPEPEEVNSVEGNSUSITCYTPPTSVNHRHTKRYKWCROGARC	62
Db	3	LEPLVTCLLAVFPAISIKSPITGPEPEEVNSVEGNSUSITCYTPPTSVNHRHTKRYKWCROGARC	62
Qy	63	GCITLLISSEGYSYSSKAGAGANLTNFPENGTFVYNALQSDSDSGRYKCGJGINSRGLSFD	122
Db	63	GCITLLISSEGYSYSSKAGAGANLTNFPENGTFVYNALQSDSDSGRYKCGJGINSRGLSFD	122
Qy	123	VSLEVSQGFLLNDTRKYVTVDLGRTVTINCEPTEMAOKRKSILYKQJGLYPLVLIIDSSGY	182
Db	123	VSLEVSQGFLLNDTRKYVTVDLGRTVTINCEPTEMAOKRKSILYKQJGLYPLVLIIDSSGY	182
Qy	183	VNPNYNGRIRLDIOGTGQLLFESVYNQLRLSDAGQYLCQACDSDSNSKKNADOLYLKPEP	242
Db	183	VNPNYNGRIRLDIOGTGQLLFESVYNQLRLSDAGQYLCQACDSDSNSKKNADOLYLKPEP	242
Qy	243	ELYEDLRGSYFPHCALGPEVANNAKFLCROSSGNCQVYNTLGKRAPAFEGRIILLNPQ	302
Db	243	ELYEDLRGSYFPHCALGPEVANNAKFLCROSSGNCQVYNTLGKRAPAFEGRIILLNPQ	302
Qy	303	DKDGSFSVYITGLRKEDAGRYLCGASHSGQLQEGSPIOAMOLFVNEESTIPRSPTYVKGV	362
Db	303	DKDGSFSVYITGLRKEDAGRYLCGASHSGQLQEGSPIOAMOLFVNEESTIPRSPTYVKGV	362
Qy	363	AGSSAVVLCPRYNKESKSIKYWCLMEGANGRCPLVDSBGMVAQYEGRLSLLEBPGNG	422
Db	363	AGSSAVVLCPRYNKESKSIKYWCLMEGANGRCPLVDSBGMVAQYEGRLSLLEBPGNG	422
Qy	423	TFPVLIIINQLTSDBAGYWCITNGDPLMTPTVEIKITIEGEPMLKPGVNTAVLGETLLKVP	482
Db	423	TFPVLIIINQLTSDBAGYWCITNGDPLMTPTVEIKITIEGEPMLKPGVNTAVLGETLLKVP	482
Qy	483	HFPCRSSEYKTYWCKKNNNGCOALPSOEDGFSKAFVNCDENSRLVSLTLNLYPRADGMY	542
Db	483	HFPCRSSEYKTYWCKKNNNGCOALPSOEDGFSKAFVNCDENSRLVSLTLNLYPRADGMY	542
Qy	543	WCVGKQGHFYGETTAAVVAVEBRKAGSRDVS LAKADAAPDEKVLDSGFREIENKA1ODP	602
Db	543	WCVGKQGHFYGETTAAVVAVEBRKAGSRDVS LAKADAAPDEKVLDSGFREIENKA1ODP	602

QY	603	RLFAE	607
		11111	
Db	603	RLFAE	607

RESULT 8
AAV73981
ID AAV73981 standard; Protein; 272 AA.

AC	AAV73981;
XX	
DT	14-MAR-2000 (first entry)

DE Human prostate tumor EST fragment derived protein #168.

KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;

XX

XX

[illegible]

FD 04 NOV 1955
XX

20-APR-1956, 38DE-1020150-
FF
XX

PR 28-APR-1998; 98DE-1020190-
XX

PA (MEIA-) MEIAGEN GES GENOMFORSCHUNG MBH.
XX

PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, 2022

DR WPI: 1999-621386/54

1000

PT proteins -

PS Claim 23; Page 379; 502pp; German.

CC This invention describes novel polypeptides and their encoding nucleic
CC acids derived from human pancreatic tumor tissue which have cytostatic
CC activity. The sequences are also useful in producing pharmaceutical
CC compositions for treatment of pancreatic tumors. AA173814-174252
CC represent protein treatments encoded by the human pancreatic tumor cDNA
CC library derived expressed sequence tag (EST) sequences represented in
CC AA52858-253014.

SQ Sequence 272 AA;

```

Query Match      33.2%; Score 254; DB 20; Length 272;
Best Local Similarity 100.0%; Pred. No. 2.7e-251;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

```

QY	511	EGPSKAFCNCGENSLVLTJNLNLRADBGWMCCKOGHFGEPAAYVVAEERKAACS	570
Db	19	EGPSAFAPNCGENSLVLTJNLNLRADBGWMCCKOGHFGETAAYVVAEERKAACS	78
QY	571	RDVSLAKADAAPEDEKVLDSGPREIENKKAIODEPLAEAEKXAVADTRDQADGSRAVDSGSS	630
Db	79	RDVSLAKADAAPEDEKVLDSGPREIENKKAIODEPLAEAEKXAVADTRDQADGSRAVDSGSS	138
QY	631	EEGGSSRALVSTLCVPLGLVLYAVGAVGAARHRRKNVDVRSISRYPTDIMSDEPNR	690
Db	139	EEGGSSRALVSTLCVPLGLVLYAVGAVGAARHRRKNVDVRSISRYPTDIMSDEPNR	198
QY	691	EEGANDNNKASSITQETSLGCKEEFVATTESTTEKKEPKKARRSSKEEEMAYKDFLLDS	750
Db	199	EEGANDNNKASSITQETSLGCKEEFVATTESTTEKKEPKKARRSSKEEEMAYKDFLLDS	258
QY	751	STVAEAODGPOEA 764	

Db 259 STVAEAKDQPOEA 272

RESULT 9

ABP55307
ID ABP55307 standard; Protein: 243 AA.

AC ABP55307;

DT 28-JAN-2003 (first entry)

DE Human polyimmunoglobulin receptor (pIgR) amino acid sequence.

KM Transseptinial transport; membrane bound vesicle; virion; liposome;
KM envelope; capsid; transmembrane domain; gene therapy; immunostimulant;
KM cytosolic; haemostatic; neuroprotective; antirheumatic; antiarthritic;
KM anticancer; antibacterial; anti-HIV; hepatotropic; virucide; exocytosis;
KM antiinflammatory; apical endocytosis; basolateral endocytosis; ADA-SCID;
KM transcytosis; monogenic disease; ADA deficiency; cystic fibrosis; ALS;
KM X-linked severe combined immunodeficiency; Haemophilia B; cancer; HIV;
KM chronic granulomatous disease; coronary artery disease; viral infection;
KM amyotrophic lateral sclerosis; rheumatoid arthritis; hepatitis; Herpes;
KM pathogenic disorder; human immunodeficiency virus; bacterial infection;
KM tuberculosis; Chlamydia; gastrointestinal ulcer; pIgR;
KM polyimmunoglobulin receptor.

OS Homo sapiens.

PN WO200283840-A2.

PD 24-OCT-2002.

PT 03-APR-2002; 2002WO-US10647.

PR 03-APR-2001; 2001US-281275P.

PA (ARIZ-) ARIZEKE PHARM INC.

PI Sheridan PL, Houston LL;

DR WPI; 2003-046923/04.

XX Fusion protein which confers the ability to penetrate epithelial cell
XX layer and to undergo paracellular transport, has a transseptinial
XX PT delivery element and a transmembrane domain from different proteins
XX PS Disclosure: Fig 2B: 160pp; English.

XX The present invention describes a fusion protein (I) comprising a
XX transseptinial delivery element (TDE) from a first protein and a
XX transmembrane domain from a second protein, or comprising TDE and a
XX viral sequence that confers the ability to be associated with or
XX incorporated into an envelope or capsid protein of a virus. (I) has
XX immunostimulant, cytosolic, haemostatic, neuroprotective, antirheumatic,
XX antiarthritic, anticancer, antibacterial, anti-HIV, hepatotropic,
XX virucide and antiinflammatory activities, and can be used in gene
XX therapy. (I) confers the ability to undergo apical endocytosis,
XX basolateral endocytosis, apical or basolateral exocytosis, apical to
XX basolateral transcytosis and basolateral to apical transcytosis. Diseases
XX treatable by gene therapy include monogenic diseases such as X-linked
XX severe combined immunodeficiency, ADA deficiency (ADA-SCID), cystic
XX fibrosis, Haemophilia B, chronic granulomatous disease, cancers such as
XX ovarian cancer, other diseases such as coronary artery disease,
XX amyotrophic lateral sclerosis (ALS), rheumatoid arthritis, pathogenic
XX disorders, including human immunodeficiency virus (HIV), viral
XX infections, hepatitis, non-specific bacterial infection, tuberculosis,
XX Herpes, Chlamydia and gastrointestinal ulcer. The present sequence
XX represents a polyimmunoglobulin receptor (pIgR) amino acid sequence
XX which is given in the exemplification of the present invention.

CC Sequence 243 AA;

Query Match

31.8%; Score 243; DB 24; Length 243;

Best Local Similarity 100.0%; Pred. No. 4.6e-240; Indels 0; Gaps 0;
Matches 243; Conservative 0; Mismatches 0;

OY 492 EKYMCKWNTGCOALPSPDEGPKAFVNCDENSRLVLTUNLYTRADEGMVCGYKQGHF 551
DB 1 EKYMCKWNTGCOALPSPDEGPKAFVNCDENSRLVLTUNLYTRADEGMVCGYKQGHF 60
OY 552 YGETAAVYVAVEERKAAGSRDYSKADAAPEEKVLDGSGFREIENKAIDPRLFAEEKAV 611
DB 61 YGETAAVYVAVEERKAAGSRDYSKADAAPEEKVLDGSGFREIENKAIDPRLFAEEKAV 120
OY 612 ADTRODADSRASVDSGSSSEEGGSSRALVSTPLGLVANGAVGARRRHKKNDR 671
DB 121 ADTRODADSRASVDSGSSSEEGGSSRALVSTPLGLVANGAVGARRRHKKNDR 180
OY 672 VSTRSRYTDSMSDFNSREFGANDMGASITTOETSLGCKEFPATTESTETKEPKKA 731
DB 181 VSTRSRYTDSMSDFNSREFGANDMGASITTOETSLGCKEFPATTESTETKEPKKA 240
OY 732 KRS 734
DB 241 KRS 243

RESULT 10

ABP55315
ID ABP55315 standard; Protein: 243 AA.

AC ABP55315;

DT 28-JAN-2003 (first entry)

DE Human polyimmunoglobulin receptor (pIgR) amino acid sequence.

KM Transseptinial transport; membrane bound vesicle; virion; liposome;
KM envelope; capsid; transmembrane domain; gene therapy; immunostimulant;
KM cytosolic; haemostatic; neuroprotective; antirheumatic; antiarthritic;
KM anticancer; antibacterial; anti-HIV; hepatotropic; virucide; exocytosis;
KM antiinflammatory; apical endocytosis; basolateral endocytosis; ADA-SCID;
KM transcytosis; monogenic disease; ADA deficiency; cystic fibrosis; ALS;
KM X-linked severe combined immunodeficiency; Haemophilia B; cancer; HIV;
KM chronic granulomatous disease; coronary artery disease; viral infection;
KM amyotrophic lateral sclerosis; rheumatoid arthritis; hepatitis; Herpes;
KM pathogenic disorder; human immunodeficiency virus; bacterial infection;
KM tuberculosis; Chlamydia; gastrointestinal ulcer; pIgR;
KM polyimmunoglobulin receptor.

OS Homo sapiens.

PN WO200283840-A2.

PD 24-OCT-2002.

PT 03-APR-2002; 2002WO-US10647.

PR 03-APR-2001; 2001US-281275P.

PA (ARIZ-) ARIZEKE PHARM INC.

PI Sheridan PL, Houston LL;

DR WPI; 2003-046923/04.

XX Fusion protein which confers the ability to penetrate epithelial cell
XX layer and to undergo paracellular transport, has a transseptinial
XX PT delivery element and a transmembrane domain from different proteins
XX PS Disclosure: Fig 2D: 160pp; English.

XX The present invention describes a fusion protein (I) comprising a
XX transseptinial delivery element (TDE) from a first protein and a
XX transmembrane domain from a second protein, or comprising TDE and a
XX viral sequence that confers the ability to be associated with or

CC incorporated into an envelope or capsid protein of a virus. (1) has
 CC immunostimulant, cytostatic, haemostatic, neuroprotective, antirheumatic,
 CC antitachytic, antitumor, antibacterial, anti-HIV, hepatotropic,
 CC antiviral and antiinflammatory activities, and can be used in gene
 CC therapy. (1) confers the ability to undergo apical endocytosis.
 CC basolateral endocytosis, apical or basolateral exocytosis, apical to
 CC basolateral transcytosis and basolateral to apical transcytosis. Diseases
 CC treatable by gene therapy include monogenic diseases such as X-linked
 CC severe combined immunodeficiency, ADA deficiency (ADA-SCID), cystic
 CC fibrosis, Haemophilia B, chronic granulomatous disease, cancers such as
 CC ovarian cancer, other diseases such as coronary artery disease,
 CC amyotrophic lateral sclerosis (ALS), rheumatoid arthritis, pathogenic
 CC disorders, including human immunodeficiency virus (HIV), viral
 CC infections, hepatitis, non-specific bacterial infection, tuberculosis,
 CC Herpes, Chlamydia and gastrointestinal ulcer. The present sequence
 CC represents a polyimmunoglobulin receptor (PIgR) amino acid sequence
 CC which is given in the exemplification of the present invention.

XX Sequence 243 AA:

Query Match 31.8%; Score 243; DB 24; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4,6e-240;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 492 EKWKCKNNMTGQALPSQDGPSPKAFVNCDBNSRLVSLTNLVTTRADEGMYCCKGKH 551
 DB 1 EKWKCKNNMTGQALPSQDGPSPKAFVNCDBNSRLVSLTNLVTTRADEGMYCCKGKH 60
 OY 552 YGTAAYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIODPRLFAEKKAV 611
 DB 61 YGTAAYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIODPRLFAEKKAV 120
 OY 612 ADTRDQADGSRASVDSGSSSEEGSSRALVSTVPLGLVAVGAVAVGARRRRKKNVDR 671
 DB 121 ADTRDQADGSRASVDSGSSSEEGSSRALVSTVPLGLVAVGAVAVGARRRRKKNVDR 180
 OY 672 VTSRSTRTDIMSDFENSRFGANDNMGASITDFTSLGKKEEVAATTESTTEKPKKA 731
 DB 181 VTSRSTRTDIMSDFENSRFGANDNMGASITDFTSLGKKEEVAATTESTTEKPKKA 240
 OY 732 KRS 734
 DB 241 KRS 243

RESULT 11

ABP5311 ID ABP5311 standard; Protein: 94 AA.

AC ABP5311:

DT 28-JAN-2003 (first entry)

XX Human polyimmunoglobulin receptor (PIgR) stalk region.

XX Transmembrane transport; membrane bound vesicle; viron; liposome;
 KW envelope; capsid; transmembrane domain; gene therapy; immunostimulant;
 KW cytosolic; haemostatic; neuroprotective; antirheumatic; antitachytic;
 KW antitumor; antibacterial; anti-HIV; hepatotropic; antiviral; exocytosis;
 KW antiinflammatory; apical endocytosis; basolateral endocytosis; ADA-SCID;
 KW transcytosis; monogenic disease; ADA deficiency; cystic fibrosis; ALS;
 KW X-linked severe combined immunodeficiency; Haemophilia B; cancer; HIV;
 KW chronic granulomatous disease; coronary artery disease; viral infection;
 KW amyotrophic lateral sclerosis; rheumatoid arthritis; hepatitis; Herpes;
 KW pathogenic disorder; human immunodeficiency virus; bacterial infection;
 KW tuberculosis; Chlamydia; gastrointestinal ulcer; PIgR;
 KW polyimmunoglobulin receptor.

XX Homo sapiens.

XX WO200283840-A2.

XX 24-OCT-2002.

XX 03-APR-2002; 2002WO-US10647.
 XX 03-APR-2001; 2001US-281275P.
 XX (ARI2-) ARIEKE PHARM INC.
 XX Sheridan PL, Houston LL;
 DR WPI; 2003-046923/04.

PT Fusion protein which confers the ability to penetrate epithelial cell
 PT layer and to undergo paracellular transport, has a transmembrane
 PT delivery element and a transmembrane domain from different proteins
 XX Disclosure; Fig 2C; 160pp; English.

XX The present invention describes a fusion protein (1) comprising a
 CC transmembrane delivery element (TDE) from a first protein and a
 CC transmembrane domain from a second protein, or comprising TDE and a
 CC viral sequence that confers the ability to be associated with or
 CC incorporated into an envelope or capsid protein of a virus. (1) has
 CC immunostimulant, cytostatic, haemostatic, neuroprotective, antirheumatic,
 CC antitachytic, antitumor, antibacterial, anti-HIV, hepatotropic,
 CC antiviral and antiinflammatory activities, and can be used in gene
 CC therapy. (1) confers the ability to undergo apical endocytosis,
 CC basolateral endocytosis, apical or basolateral exocytosis, apical to
 CC basolateral transcytosis and basolateral to apical transcytosis. Diseases
 CC treatable by gene therapy include monogenic diseases such as X-linked
 CC severe combined immunodeficiency, ADA deficiency (ADA-SCID), cystic
 CC fibrosis, Haemophilia B, chronic granulomatous disease, cancers such as
 CC ovarian cancer, other diseases such as coronary artery disease,
 CC amyotrophic lateral sclerosis (ALS), rheumatoid arthritis, pathogenic
 CC disorders, including human immunodeficiency virus (HIV), viral
 CC infections, hepatitis, non-specific bacterial infection, tuberculosis,
 CC Herpes, Chlamydia and gastrointestinal ulcer. The present sequence
 CC represents a polyimmunoglobulin receptor (PIgR) amino acid sequence
 CC which is given in the exemplification of the present invention.

SO Sequence 94 AA:

Query Match 12.3%; Score 94; DB 24; Length 94;
 Best Local Similarity 100.0%; Pred. No. 1.1e-87;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 545 GVKQGHFYGETAAYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIODPRL 604

DB 1 GVKQGHFYGETAAYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIODPRL 60

OY 605 FAEKAVADTRDQADGSRASVDSGSSSEEGGSSR 638

DB 61 FAEKAVADTRDQADGSRASVDSGSSSEEGGSSR 94

RESULT 12

AAG65712 ID AAG65712 standard; protein: 90 AA.

AC AAG65712:

DT 07-JAN-2002 (first entry)

XX Human polymeric immunoglobulin receptor (PIgR) fragment.

XX Polymeric immunoglobulin receptor; PIgR; ligand; therapeutic;
 KW carcinoma diagnosis; veterinary; human.

XX Homo sapiens.

XX WO200172846-A2.

XX 04-OCT-2001.

PF 26-MAR-2001: 2001MO-US09699.
 XX
 PR 27-MAR-2000: 2000US-192197P.
 PR 27-MAR-2000: 2000US-192198P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Mostov KE, Chapin SJ, Richman-Eisenstat J;
 XX
 DR WPI; 2001-611619/70.
 XX
 PT New ligands binding to a specific region of a polymetric immunoglobulin
 PT receptor, useful for transporting therapeutic or diagnostic
 PT compositions into or across cells expressing pigR e.g. in drug delivery
 PT
 XX
 PS Disclosure: Fig 3: 102pp: English.
 XX
 CC The invention provides ligands that bind specifically to a region of an
 CC animal cell polymetric immunoglobulin receptor (pigR). The pigR cleaves
 CC to produce a stalk region remaining attached to the cell and a secretory
 CC component existing in the organ of interest in several forms. The ligands
 CC do not bind to the stalk or the most abundant form of the secretory
 CC component present in the organ under physiological conditions. The
 CC ligands are useful for transporting therapeutic or diagnostic
 CC compositions into or across cells expressing pigR, useful to introduce
 CC or transport ligands such as antibodies and/or to deliver biologically
 CC active components such as proteins, nucleic acids or detectable labels.
 CC They are used to deliver therapeutic compositions to mucosal surfaces.
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.
 CC They are also useful to label cells expressing pigR, e.g. to distinguish
 CC epithelial cells from a mixed cell population in pathology studies or to
 CC aid in carcinoma diagnosis (since pigR expression is reduced in
 CC carcinomas relative to normal epithelium). They can also be used to
 CC deliver veterinary compositions, especially in mammals such as farm,
 CC domestic or wild mammals or birds e.g. birds reared for human
 CC consumption. The present sequence represents a human pigR fragment.
 CC
 XX
 SO Sequence 90 AA:

 Query Match 11.8%; Score 90; DB 22; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.3e-83;
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 549 GHFYGETAAVYVVEERKAAGSRDYS LAKADAAPEKVLDSGFREIENKAIDPRLFAEE 608
 DB 1 GHFYGETAAVYVVEERKAAGSRDYS LAKADAAPEKVLDSGFREIENKAIDPRLFAEE 60
 QY 609 KAVADTRDQADGSRASVDSGSSSEEGGSSR 638
 DB 61 KAVADTRDQADGSRASVDSGSSSEEGGSSR 90

 RESULT 13
 AAM43098
 ID AAM43098 standard; peptide; 61 AA.
 AC AAM43098;
 XX
 DT 04-JUN-1998 (first entry)
 XX
 DE Polymetric immunoglobulin receptor (pigR) stalk sequence 1.
 XX
 KW Polymetric immunoglobulin receptor; pigR; stalk; epithelial cell;
 KW ligand; antibody; target; binding; mammalian.
 XX
 OS Mammalia.
 XX
 PN WO9746588-A1.
 XX
 PD 11-DEC-1997.
 XX
 PR 14-MAY-1997; 97WO-US07944.
 PF

XX
 PR 04-JUN-1996; 96US-0018958.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Mostov KE, Richman-Eisenstat J;
 XX
 DR WPI; 1998-042123/04.
 XX
 PT ligand that binds the stalk of a cell's polymetric immunoglobulin
 PT receptor - useful to target to, into or across mammalian epithelial
 PT cell biologically active component, e.g. nucleic acid, protein,
 PT lipid, carbohydrate, etc
 PT
 XX
 PS Claim 28; Page 37; 42pp: English.
 XX
 CC This peptide sequence represents the stalk of the polymetric
 CC immunoglobulin receptor (pigR) to which a ligand can bind to. The stalk
 CC is the extracellular component of the pigR that is bound to the cell
 CC following cleavage of the secretory component of the pigR. The stalk is
 CC present regardless of whether the secretory component segment is cleaved
 CC or uncleaved from pigR. A ligand, preferably a humanised antibody or a
 CC recombinant single chain variable region fragment can specifically bind
 CC to the stalk of a pigR of a cell under physiological conditions, but not
 CC to the secretory component of pigR. Such a ligand can be introduced into
 CC a cell expressing a pigR by attaching to the stalk of the pigR. The
 CC ligand can be used to target to, into or across the apical or basolateral
 CC surface of a mammalian epithelial cell, a biologically active component
 CC selected from a nucleic acid (preferably encoding the wild type cystic
 CC fibrosis transmembrane conductance regulator), protein, radioisotope,
 CC lipid or carbohydrate. The biologically active composition can also be
 CC selected from a group consisting of anti-inflammatories, antisense
 CC oligonucleotides, antibiotics or anti-infectives.
 CC
 XX
 SO Sequence 61 AA:

 Query Match 8.0%; Score 61; DB 19; Length 61;
 Best Local Similarity 100.0%; Pred. No. 4.9e-54;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 578 ADAAPDEKVLDSGFREIENKAIDPRLFAEEKAVADTRDQADGSRASVDSGSSSEEGGSS 637
 DB 1 ADAAPDEKVLDSGFREIENKAIDPRLFAEEKAVADTRDQADGSRASVDSGSSSEEGGSS 60
 QY 638 R 638
 DB 61 R 61

 RESULT 14
 AAM43099
 ID AAM43099 standard; peptide; 61 AA.
 AC AAM43099;
 XX
 DT 04-JUN-1998 (first entry)
 XX
 DE Polymetric immunoglobulin receptor (pigR) stalk sequence 2.
 XX
 KW Polymetric immunoglobulin receptor; pigR; stalk; epithelial cell;
 KW ligand; antibody; target; binding; mammalian.
 XX
 OS Mammalia.
 XX
 PN WO9746588-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 14-MAY-1997; 97WO-US07944.
 XX
 PR 04-JUN-1996; 96US-0018958.
 XX
 PA (REGC) UNIV CALIFORNIA.

```

XX  PI      Mostov KE, Richman-Eisenstat J;
XX  DR      WPI: 1998-042123/04.
XX  XX
XX  PT      Ligand that binds the stalk of a cell's polymeric immunoglobulin
XX  PT      receptor - useful to target to, into or across mammalian epithelial
XX  PT      cell biologically active component, e.g. nucleic acid, protein,
XX  PT      lipid, carbohydrate, etc
XX  PS      Claim 28; Page 37; 42pp; English.
XX  CC
XX  CC      This peptide sequence represents the stalk of the polymeric
XX  CC      immunoglobulin receptor (PIGR) to which a ligand can bind to. The stalk
XX  CC      is the extracellular component of the PIGR that is bound to the cell
XX  CC      following cleavage of the secretory component of the PIGR. The stalk is
XX  CC      present regardless of whether the secretory component segment is cleaved
XX  CC      or uncleaved from PIGR. A ligand, preferably a humanised antibody or a
XX  CC      recombinant single chain variable region fragment can specifically bind
XX  CC      to the stalk of a PIGR of a cell under physiological conditions, but not
XX  CC      to the secretory component of PIGR. Such a ligand can be introduced into
XX  CC      a cell expressing a PIGR by attaching to the stalk of the PIGR. The
XX  CC      ligand can be used to target to, into or across the apical or basolateral
XX  CC      surface of a mammalian epithelial cell, a biologically active component
XX  CC      selected from a nucleic acid (preferably encoding the wild type cystic
XX  CC      fibrosis transmembrane conductance regulator), protein, radioisotope,
XX  CC      lipid or carbohydrate. The biologically active composition can also be
XX  CC      selected from a group consisting of anti-inflammatory, antisense
XX  CC      oligonucleotides, antibiotics or anti-infectives.
XX  SQ      Sequence 61 AA:
SQ
Query Match      8.0%; Score 61; DB 19; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.9e-54;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      578 ADAAPDEKVLDSGFREIENKAIDPRLFAEKKAVADTRDQADGSRASVDSGSSSEEGGSS 637
DB      1 ADAAPDEKVLDSGFREIENKAIDPRLFAEKKAVADTRDQADGSRASVDSGSSSEEGGSS 60
OY      638 R 638
DB      61 R 61

```

RESULT 15
AAW99073
ID AAW99073 standard; Protein: 60 AA.
XX
AC AAW99073;
XX
DT 13-MAY-1999 (first entry)
XX
DE Immunoglobulin like protein PolyIgrV4.
XX
KW Human; PIGR-1; immunoglobulin; rheumatoid arthritis; psoriasis;
KW multiple sclerosis; systemic lupus erythematosus; diagnosis; receptor;
KW inflammatory bowel disease.
XX
OS Unidentified.
XX
PN EP897981-A1.
PD 24-FEB-1999.
XX
PE 11-AUG-1998; 98EP-0306403.
XX
PR 22-OCT-1997; 97US-0955937.
PR 19-AUG-1997; 97US-0056152.
XX
PA (SMIK) SMITHKLINE BEECHAM.
XX
PI Hurtle MR, Sweet RW, Truneh A, Wu S;

```

XX  DR      WPI: 1999-134644/12.
XX  XX
XX  PT      New receptor (PIGR-1) polypeptide and polynucleotide - useful as
XX  PT      diagnostic reagents and for prevention and treatment of multiple
XX  PT      sclerosis, inflammatory bowel disease and psoriasis
XX  PS      Example 2; Page 14; 28pp; English.
XX  CC
XX  CC      The present invention describes a new receptor polypeptide designated
XX  CC      PIGR-1, which is a member of the immunoglobulin (Ig) superfamily. PIGR-1
XX  CC      proteins and polynucleotides are useful for diagnosing susceptibility to
XX  CC      diseases by detecting mutations in the PIGR-1 gene, and can diagnose
XX  CC      diseases associated with PIGR-1 protein imbalance by determining PIGR-1
XX  CC      protein expression levels. PIGR-1 proteins can be used to screen for
XX  CC      agonists and antagonists by measuring the binding to protein, and
XX  CC      observing the protein function. These can be used in treatment to
XX  CC      activate (agonist) or inhibit (antagonist) PIGR-1 activity. In addition
XX  CC      to direct administration of antisense sequences to prevent expression, or
XX  CC      PIGR-1 polynucleotides to treat conditions associated with a lack of
XX  CC      PIGR-1 protein. Gene therapy may also be used to affect endogenous PIGR-1
XX  CC      protein expression. PIGR-1 antibodies are useful for inducing an immune
XX  CC      response to immunise and prevent disease, and for isolating PIGR-1
XX  CC      clones or purifying the polypeptides by affinity chromatography. PIGR-1
XX  CC      proteins can be administered directly or as a vaccine and treated include:
XX  CC      against disease. Diseases diagnosed, prevented and treated include:
XX  CC      rheumatoid arthritis, multiple sclerosis, psoriasis, systemic lupus
XX  CC      erythematosus and inflammatory bowel disease. The PIGR-1 protein is also
XX  CC      useful for mapping the gene to a chromosome, allowing gene inheritance
XX  CC      to be studied through linkage analysis. The present sequence represents
XX  CC      an immunoglobulin like protein from the present invention.
XX  SQ      Sequence 60 AA:
SQ
Query Match      7.9%; Score 60; DB 20; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      353 PRSPVTYKGVAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPPLVDSGQVKAQYEGR 412
DB      1 PRSPVTYKGVAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPPLVDSGQVKAQYEGR 60

```

Search completed: September 9, 2003, 23:17:14
Job time : 86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:15:47 ; Search time 32 Seconds
(without alignments)
1010.172 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764
Sequence: 1 MLFLVLTCLAVPAISTNS.....DFLLQSTVAEAADGPGQA 764

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents_Aa: *
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/1/1aa/5A.COMB.pep: *
4: /cgn2_6/prodata/1/1aa/5B.COMB.pep: *
5: /cgn2_6/prodata/1/1aa/5C.COMB.pep: *
6: /cgn2_6/prodata/1/1aa/5D.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	746	97.6	746	3 US-08-434-000A-4	Sequence 4, Appl 1
2	746	97.6	746	4 US-09-312-157-4	Sequence 4, Appl 1
3	607	79.5	608	4 US-09-095-385-4	Sequence 4, Appl 1
4	61	8.0	61	3 US-08-856-383-10	Sequence 10, Appl 1
5	61	8.0	61	3 US-08-856-383-11	Sequence 11, Appl 1
6	61	8.0	61	4 US-09-475-088-10	Sequence 10, Appl 1
7	61	8.0	61	4 US-09-475-088-11	Sequence 11, Appl 1
8	60	7.9	60	3 US-08-955-937A-6	Sequence 6, Appl 1
9	60	7.9	60	3 US-09-300-985-6	Sequence 6, Appl 1
10	53	6.9	109	3 US-08-961-564A-9	Sequence 9, Appl 1
11	43	5.6	43	3 US-08-955-937A-11	Sequence 11, Appl 1
12	43	5.6	43	3 US-09-300-985-11	Sequence 11, Appl 1
13	31	4.1	31	3 US-08-856-383-4	Sequence 4, Appl 1
14	31	4.1	31	4 US-09-475-088-4	Sequence 4, Appl 1
15	29	3.8	46	3 US-08-955-937A-10	Sequence 10, Appl 1
16	29	3.8	46	3 US-09-300-985-10	Sequence 10, Appl 1
17	27	3.5	57	3 US-08-955-937A-5	Sequence 5, Appl 1
18	27	3.5	57	3 US-09-300-985-5	Sequence 5, Appl 1
19	21	2.7	757	4 US-08-434-000A-6	Sequence 6, Appl 1
20	21	2.7	757	4 US-09-312-157-6	Sequence 6, Appl 1
21	19	2.5	40	3 US-08-856-383-6	Sequence 6, Appl 1
22	19	2.5	40	4 US-09-475-088-6	Sequence 6, Appl 1
23	17	2.2	769	3 US-08-434-000A-10	Sequence 10, Appl 1
24	17	2.2	769	3 US-09-312-157-10	Sequence 10, Appl 1
25	16	2.1	16	3 US-08-782-480-45	Sequence 45, Appl 1
26	16	2.1	16	3 US-08-954-211-45	Sequence 45, Appl 1
27	16	2.1	16	4 US-09-005-167A-45	Sequence 45, Appl 1

28	16	2.1	16	4 US-09-176-741B-45	Sequence 45, Appl 1
29	16	2.1	771	3 US-08-434-000A-8	Sequence 8, Appl 1
30	16	2.1	771	4 US-09-312-157-8	Sequence 8, Appl 1
31	13	1.7	773	3 US-08-434-000A-2	Sequence 2, Appl 1
32	13	1.7	773	4 US-09-312-157-2	Sequence 2, Appl 1
33	12	1.6	16	2 US-08-656-906-1	Sequence 1, Appl 1
34	12	1.6	16	3 US-09-217-847-1	Sequence 1, Appl 1
35	9	1.2	624	2 US-08-642-406A-22	Sequence 22, Appl 1
36	9	1.2	624	4 US-09-199-534-22	Sequence 22, Appl 1
37	9	1.2	60	3 US-08-961-564A-6	Sequence 6, Appl 1
38	8	1.0	417	4 US-09-252-991B-27213	Sequence 27213, A
39	8	1.0	777	2 US-08-874-678-3	Sequence 3, Appl 1
40	8	1.0	777	3 US-08-643-839-3	Sequence 3, Appl 1
41	8	1.0	777	4 US-09-348-886-3	Sequence 3, Appl 1
42	8	1.0	802	4 US-09-173-151A-33	Sequence 33, Appl 1
43	8	1.0	1298	1 US-08-222-616-33	Sequence 33, Appl 1
44	8	1.0	1298	1 US-08-340-011-2	Sequence 2, Appl 1
45	8	1.0	1298	1 US-08-340-011-2	Sequence 2, Appl 1

ALIGNMENTS

RESULT 1
US-08-434-000A-4
Sequence 4, Application US/08434000A
Patent No. 6046037
GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA: US/08/434,000A
APPLICATION NUMBER: US/08/434,000A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/367,395
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
? HUMAN POLYIMMUNOGLOBULIN RECEPTOR
US-08-434-000A-4

Query Match 97.6%; Score 746; DB 3; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 78
1 KSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 60
79 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDRK 138
61 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDRK 120
139 VYTVDLGRVTITNCPEKTEAOKRSKLYKQIGLYPLVYDSSGVNPNYTRGLDIOGT 198
121 VYTVDLGRVTITNCPEKTEAOKRSKLYKQIGLYPLVYDSSGVNPNYTRGLDIOGT 180
199 GOLLEFSVINQLRLSDAGGYLCQAGDSSNKKNNADLYLKPEPELYEDLRGSVTFHCA 258
181 GOLLEFSVINQLRLSDAGGYLCQAGDSSNKKNNADLYLKPEPELYEDLRGSVTFHCA 240
259 LGPEVANNAKFLCROSSGGENCDVYVNTLGKRAFAFEGRIILNPQDKDGSFVITGLRKE 318
241 LGPEVANNAKFLCROSSGGENCDVYVNTLGKRAFAFEGRIILNPQDKDGSFVITGLRKE 300
319 DAGRYLCGAHSDGLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSVAVLCPPYRKES 378
301 DAGRYLCGAHSDGLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSVAVLCPPYRKES 360
379 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLEBPGNGTFTVILNQLTSRDAGF 438
361 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLEBPGNGTFTVILNQLTSRDAGF 420
439 YWCLTNGDTLMRTTVEIKIIEGEBNLKVPGNNTAVLGETLKVPCHPCKSSYEKYWKW 498
421 YWCLTNGDTLMRTTVEIKIIEGEBNLKVPGNNTAVLGETLKVPCHPCKSSYEKYWKW 480
499 NNTGCOALPSQDEGSKAFVNCDENSRLVSLTLNLTVRADGKWCVCVKGHFGEYTAAY 558
481 NNTGCOALPSQDEGSKAFVNCDENSRLVSLTLNLTVRADGKWCVCVKGHFGEYTAAY 540
559 YVAVEERKAASRDVSLAKADAAPEDEVLDSGFEIEBNKAIDPRLEAEKAAVADTRQA 618
541 YVAVEERKAASRDVSLAKADAAPEDEVLDSGFEIEBNKAIDPRLEAEKAAVADTRQA 600
619 DGSRAVDSGSSSEBOGSSRALYSLVPLGLVLAAGVAVARARRKRVDRYSISYR 678
601 DGSRAVDSGSSSEBOGSSRALYSLVPLGLVLAAGVAVARARRKRVDRYSISYR 660
679 TDISMDFENSREPGANDNMGASITQETSLGKEEVPATTESTETKEPKKARSSKEE 738
661 TDISMDFENSREPGANDNMGASITQETSLGKEEVPATTESTETKEPKKARSSKEE 720
739 AEMAYKDFLLQSSSTVAEAADGPOEA 764
721 AEMAYKDFLLQSSSTVAEAADGPOEA 746

RESULT 2
US-09-312-157-4
Sequence 4, Application US/09312157
Patent No. 6303341

GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles

STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,157
FILING DATE: 14-May-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/434,000
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-351

SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

DESCRIPTION: Human Polymunoglobulin Receptor
SEQUENCE DESCRIPTION: SEQ ID NO: 4
US-09-312-157-4

Query Match 97.6%; Score 746; DB 4; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 78
1 KSPIFGPEEVNSVEGNSVITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 60
79 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDRK 138
61 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDRK 120
139 VYTVDLGRVTITNCPEKTEAOKRSKLYKQIGLYPLVYDSSGVNPNYTRGLDIOGT 198
121 VYTVDLGRVTITNCPEKTEAOKRSKLYKQIGLYPLVYDSSGVNPNYTRGLDIOGT 180
199 GOLLEFSVINQLRLSDAGGYLCQAGDSSNKKNNADLYLKPEPELYEDLRGSVTFHCA 258
181 GOLLEFSVINQLRLSDAGGYLCQAGDSSNKKNNADLYLKPEPELYEDLRGSVTFHCA 240
259 LGPEVANNAKFLCROSSGGENCDVYVNTLGKRAFAFEGRIILNPQDKDGSFVITGLRKE 318
241 LGPEVANNAKFLCROSSGGENCDVYVNTLGKRAFAFEGRIILNPQDKDGSFVITGLRKE 300
319 DAGRYLCGAHSDGLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSVAVLCPPYRKES 378
301 DAGRYLCGAHSDGLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSVAVLCPPYRKES 360
379 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLEBPGNGTFTVILNQLTSRDAGF 438
361 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLEBPGNGTFTVILNQLTSRDAGF 420
439 YWCLTNGDTLMRTTVEIKIIEGEBNLKVPGNNTAVLGETLKVPCHPCKSSYEKYWKW 498
421 YWCLTNGDTLMRTTVEIKIIEGEBNLKVPGNNTAVLGETLKVPCHPCKSSYEKYWKW 480
499 NNTGCOALPSQDEGSKAFVNCDENSRLVSLTLNLTVRADGKWCVCVKGHFGEYTAAY 558